Applications and Industry®

March 1961

Transactions Papers

NUMBER 53

Published Bimonthly by

AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS

Communication and Electronics—March 1961

Science and Electronics Division

Communication Division

(See inside back cover)

Note to Librarians. The six bimonthly issues of "Applications and Industry," March 1961–January 1962, will also be available in a single volume (no. 80) entitled "AIEE Transactions—Part II. Applications and Industry," which includes all technical papers on that subject presented during 1961. Bibliographic references to Applications and Industry and to Part II of the Transactions are therefore equivalent.

Applications and Industry. Published bimonthly by the American Institute of Electrical Engineers, from 20th and Northampton Streets, Easton, Pa. AIEE Headquarters: 33 West 39th Street, New York 18, N. Y. Address changes must be received at AIEE Headquarters by the first of the month to be effective with the succeeding issue. Copies undelivered because of incorrect address cannot be replaced without charge. Editorial and Advertising offices: 33 West 39th Street, New York 18, N. Y. Nonmember subscription $8.00 per year (plus 75 cents extra for foreign postage payable in advance in New York exchange). Member subscriptions: one subscription at $5.00 per year to any one of three divisional publications: Communication and Electronics, Applications and Industry, or Power Apparatus and Systems: additional annual subscriptions $8.00 each. Single copies when available $1.50 each. Second-class mail privileges authorized at Easton, Pa. This publication is authorized to be mailed at the special rates of postage prescribed by Section 132.122.

Printed in United States of America

Number of copies of this issue 5,000

An Investigation of the Transient Behavior of an Aluminum Potline Installation

ALLAN GREENWOOD
MEMBER AIEE

W. C. KOTHEIMER
ASSOCIATE MEMBER AIEE

C. A. LANGLOIS
ASSOCIATE MEMBER AIEE

OVER A PERIOD of several years, a number of transformers and induction regulators supplying energy to aluminum potlines and similar installations have been damaged by insulation failures; the failures were apparently due to abnormal overvoltages. The transformers used for these applications usually step down from 13.8 kv to a secondary voltage in the range 600 to 900 volts depending upon the installation. A basic impulse level of at least 40 kv is normal for the secondary winding; insulation failures occurring on these windings are an indication of the severity of the voltage transients. An excellent photograph showing typical coil damage appears in a paper by Pope, et al.[1]

Naturally, these overvoltages have engendered considerable speculation as to their origin and, although the principal object of the work to be described here is not to determine the precise cause of overvoltages, the authors' speculations are added, as a help in understanding how the work was planned. Later, we will look at these speculations again in the light of the results obtained.

It is clear that voltage transients must be generated within the system or else enter as surges from the high-voltage line. Many indications, at least in some instances, point to the unlikelihood of the latter possibility. Bourque and Coxe have described, in unpublished work, an installation in which the transformers were supplied on their high-voltage side, through considerable lengths of underground cable. This certainly minimizes the chance of damaging surges approaching by that route. Also, potline installations frequently comprise a number of parallel transformers which share the energy of impinging surges, and it is difficult to see how transient voltages of the order experienced could be transferred to the low-voltage winding, except perhaps as an initial spike by electrostatic coupling, without flashover occurring on the high-voltage side. For these reasons our attention was focussed mainly on the possibilities of surges being developed within the system.

When one looks at a network, such as a potline-rectifier-transformer system, it would seem that the only way in which abnormally high voltages could be produced internally would be by some form of current suppression. Energy stored in the magnetic field of the circuit would thereby be released and transferred to the electric field, where it would enhance the voltage across the system capacitance. There appear to be four active elements in the network that might bring this about: the anode and cathode circuit breakers, the rectifiers, and the potline itself. Several investigators[2,3] have ruled out the circuit breakers since significant current chopping is not a characteristic of the breakers used. The authors concur in this opinion; if it is true, it seems likely that the disturbances originate in the rectifiers or the potline by some mechanism akin to chopping. Several such possible mechanisms have been proposed by Hull and Elder[4] and more recently by Steiner and Strecker,[5] and under certain conditions surging in rectifiers has been observed in controlled test. Potline cells are also known to behave rather erratically from time to time. However, in our investigation we were not concerned with mechanisms or, indeed, with fixing responsibilities on one element or another. Rather, this study starts with the assumption that abrupt changes of current could occur, and proceeds to analyze the consequences thereof.

For a better appreciation of the results of this investigation a brief résumé of the behavior of a circuit when the current in it suddenly changes, would seem to be in order. In the limit when the current changes instantaneously, the voltage generated is given by the product of the instantaneous change of current and the surge impedance of the circuit. It is not the product Ldi/dt, which in this instance would be infinite, for this ignores the capacitance of the component or circuit involved. Most practical circuits comprise a number of elements in series/parallel combinations, but on analysis they can usually be grouped into several series units each with its own surge impedance.

When such a circuit is excited by a step function of current, or by what for convenience we will call a chop of current, each unit will respond in the manner described above, i.e., it will be set into oscillation at a frequency peculiar to itself and, in the absence of damping, the peak voltage generated will be iZ_0, where Z_0 is its surge impedance. The total response of such a circuit will be the summation of the individual responses of these units, i.e., a multifrequency disturbance.

In practice it frequently happens that one part of the circuit has a much higher surge impedance than the rest so that voltage tends to pile up across this part. Again, in all practical circuits the suppression of current always takes a finite time although in some cases this may be very short indeed. Because of this the overvoltage is not as great as the theoretical limit. The discrepancy will depend upon the rate of decline of current and the character of the circuit itself. Specifically, it has been shown by Lee[6] that the overvoltage will be significant if the current collapses in a time less than half a cycle of natural frequency of the circuit involved. Declining as a linear ramp in this time, the current would generate an overvoltage of approximately 60% of the theoretical peak; for longer times, the overvoltage is progressively less. This point has a very important bearing on the results to be described below. With this background we can now outline the aims and scope of the investigation.

If the voltage transients are, indeed, caused by the sudden change of current at either a rectifier or in the potline, then by creating or simulating a chop at these locations, it should be possible to investigate the magnitude and distribution of the voltages that they give

Paper 61-20, recommended by the AIEE Industrial Power Rectifiers Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 29–February 3, 1961. Manuscript submitted October 3, 1960; made available for printing November 16, 1960.

ALLAN GREENWOOD and W. C. KOTHEIMER are with the General Electric Company, Philadelphia, Pa., and C. A. LANGLOIS is with the Reynolds Metals Company, Richmond, Va.

The authors wish to thank their colleagues, D. C. Hoffman, T. H. Lee, V. N. Stewart, and C. H. Titus, for their helpful discussions before and after the tests were carried out.

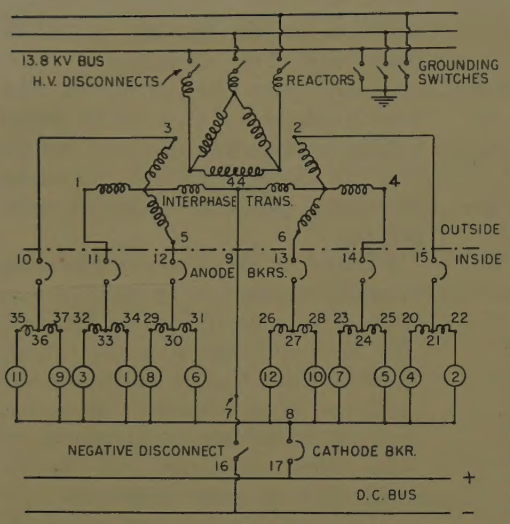


Fig. 1 (left). Circuit diagram of transformer, switchgear, and rectifiers

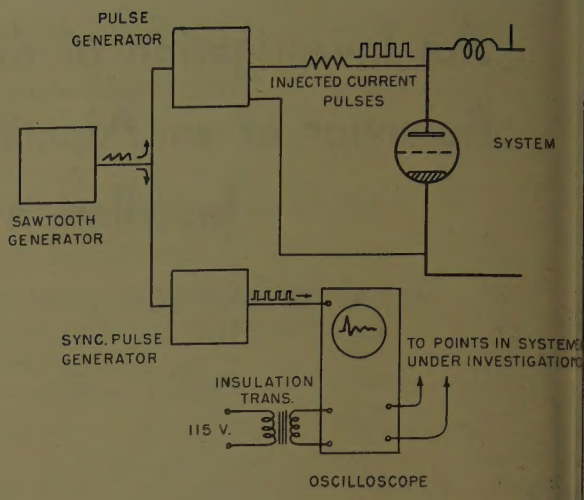


Fig. 2 (right). Test circuit for current injection technique

rise to. This is tantamount to measuring dynamically the surge impedances of the constituent parts of the circuit for various possible circuit connections, and thereby serves two purposes: (1) It permits one to determine the minimum magnitude of the current suppression that must occur to generate any specified voltage at any specified point. (2) If in any future investigation, with a surge recorder, for example, a certain pattern of transient voltages is disclosed, it should be possible by comparing the pattern with the test results to pinpoint the origin of the disturbance. It was therefore not the aim of the investigation to determine the cause of overvoltages, but rather to say that if current suppression were the cause, it must occur in a particular locality with a particular severity.

The opportunity to carry out such a test program came in June 1958, when the Reynolds Metals Company was about to commission a new aluminum potline at their Lister Hill, Ala., plant. The results obtained apply in detail to that particular system only, but in a broader way they provide information on many similar circuits. Moreover, the test methods described below can be applied equally well elsewhere.

Circuit Arrangements and Test Procedures

A schematic diagram of a typical rectifier skid and associated transformer and breaker equipment is shown in Fig. 1, where it will be noted that the transformer has two secondary windings. In fact, it had four such windings and two primaries on a common core; the others served an identical skid adjacent to the one shown. In all, there were 7 such transformers and 14 rectifier skids supplying the potline.

Two methods of testing were adopted. In the first, a direct current was fed into the system at the supposed point of suppression, e.g., at the anode and cathode of one of the rectifiers. This current, approximately 0.4 amp (ampere) was then interrupted suddenly by means of a vacuum switch. In the second method, square pulses of current were repeatedly injected into the circuit at the rectifier. The front of each square pulse represents a negative chop, as it were, while the back represents a true chop. The voltages generated by these sudden changes of current were then investigated oscillographically at all locations of interest. The circuit for the current injection method is shown in Fig. 2. It has the advantage that, because of the recurrent nature of the stimulus, the oscilloscope sweep can be synchronized with the surge generator, so that the circuit response is seen as an apparently steady picture on the screen. With the d-c interruption method the oscilloscope had to be triggered at the initiation of the transient, which then appeared as a single sweep on the screen.

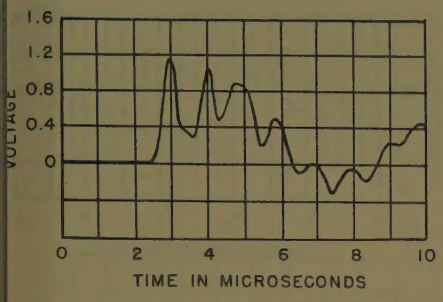
In both methods the systems were dead at the time of the tests. It was recognized that surges might occur under many different circuit conditions, viz., normal operation with or without commutation, backfire, misfire, etc.; these different conditions had to be simulated, which was a fairly simple process. Those rectifiers which were assumed to be conducting at the time the surge occurred were simply short-circuited out with clip leads; the remainder were left open. This reveals a further advantage of the recurrent surge method of testing: changes from one condition to another could be made quickly and simply while the system was being excited; the consequences of any change could be observed at once on the oscilloscope screen.

The benefits of repetitive surge technique has been very aptly summarized by Wilkinson:[7]

...a recurrent surge oscillograph is strikingly effective in its power of conveying information to the user. This is so because of the ease with which questions about the effect of any particular detail in a network under review are put to the test and immediately produce a clear fixed image on the oscillograph by way of reply. If the recurrent surge oscillograph is like a tutor with whom problems can be discussed, the single-record oscillograph is to be compared with a system of correspondence by field post card.

When the recurrent method was checked against that using the vacuum switch, the agreement was found to be remarkably good. Because of this agreement and the advantages of the current injection method, this method was used almost exclusively during the investigation. For the majority of the tests the 13.8-kv bus was short-circuited and grounded in order to simulate the normal energized condition. The source is stiff so, for all practical purposes, we can neglect its impedance. A few tests were conducted with the transformer primaries open-circuited, a condition which occurs when a potline is being de-energized by first interrupting the a-c supply. If a surge should occur in the short interval between dropping out the 13.8-kv circuit breakers and the subsequent opening of the anode and cathode breakers, then it would be simulated by the conditions just described.

So far, nothing has been said about the potline itself. It could not, of course, be introduced directly into the problem. Before the test program was started, the problem of simulating the potline was given considerable attention. The potline is physically extensive, in some ways resembling a transmission line looped back on itself; it is isolated from ground



ig. 3. Response at points 7 and 22 to 20 ma (milliamperes) injected at tank 2

n both polarities; but it has, one would uppose, considerable capacitance to round. It was felt that as a first pproximation, the simplest analog would e a parallel *L-C* (inductance-capacitance) ircuit. By temporarily disconnecting ne of the operating potlines and shock- ng it, it was hoped to establish some eliable value for the *L* and *C* compo- ents.

When the test program preliminaries ere under way, it was both interesting nd gratifying to find that the transients eing measured were relatively insensitive o the potline analog. For example, put- ing a 0.5-μf (microfarad) capacitor across l-c bus had only secondary effects on the agnitude and waveform of the voltages enerated around the rectifier skid and at he transformer, when simulated current hops were applied. The reason for this ill become apparent when some of the esults are analyzed.

Measurements

There were a number of rather unusual eatures about the system under test that resented their peculiar problems from measurement point of view. Perhaps he most important of these was the fact hat the entire system was ungrounded, hough there was capacitance coupling o ground.

Taking measurements on such a system as rather like standing on a spring attress and trying to measure the istance to the floor. This situation alled for the isolation from ground of oth the surge-generating equipment and he oscilloscope used for measurement. he equipment that generated the current ulses was fed through an isolation trans- ormer which had significantly less apacitance between its windings than ad its own power supply transformer. By increasing the impedance of this path, he errors that would be introduced by round current circulating through the ystem capacitance to ground and back o the generator through its ground capac- tance would be minim zed.

The measurement end was tackled by measuring the transient voltage differentially, i.e., a double-ended or differential amplifier was introduced between the point of measurement and the oscilloscope plates. It was apparent from oscillograms taken, that the system response was quite intimately dependent on the stray capacitance of the system. To have used a single-ended amplifier input to the oscilloscope would have been to connect the capacitance of the oscilloscope (i.e., the capacitance across its power supply to the 60-cycle system) to one of the measuring points. This would inevitably have introduced an error.

The oscillogram, Fig. 3, is typical of the kind of response that was observed. The point of interest here is the very fast rate of rise of the initial spike of voltage: it reaches its crest in approximately 0.3 μsec (μsecond). It will be recalled that at the start of the paper a special point was made of the fact that the surge generated by a current chop depends upon the rate at which the current declines; the more rapid the rate, the greater the voltage, being a maximum when the chop is vertical. To approach this maximum the time for the current to come to zero should be short compared with the period of the highest frequency excited.

In view of the obviously very high frequency that appears in Fig. 3, it was necessary to take a careful look at the front of the injected current pulse to make quite certain that it was fast enough to develop a voltage approaching the maximum. Fig. 4 is a record of the current pulse wavefront; note that the current reaches approximately 80% of the maximum value and perhaps 90% of the final value on 0.2 μsec. It is there-

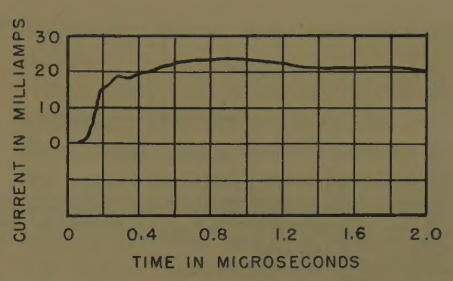


Fig. 4. Profile of injected current

fore safe to assert that the response shown in Fig. 3 is not less than 85% of the maximum available from a vertical chop.

The rise time of the initial transient in Fig. 3 raised the question as to whether the response of the amplifier was adequate. This point was checked with some care using different amplifiers. Widening the bandwidth did not change the appearance of the fast transients.

Several other smaller difficulties had to be disposed of before the test procedure ran smoothly. The physical distance involved caused some inconvenience; in many instances the point of measurement was quite remote from the point where the stimulus was applied to the circuit. For example, on one occasion, measurements were made at the bushings on top of the transformer, situated outside in the yard, and the pulsing equipment was located inside the rectifier room close to one of the skids.

Synchronizing becomes difficult under such conditions, the firing circuits of neighboring energized potlines put out a constant procession of pulses giving a high background noise level, which calls for careful screening of trigger circuit leads to the oscilloscope. It is usual to have the pulse generator initiate the oscillo-

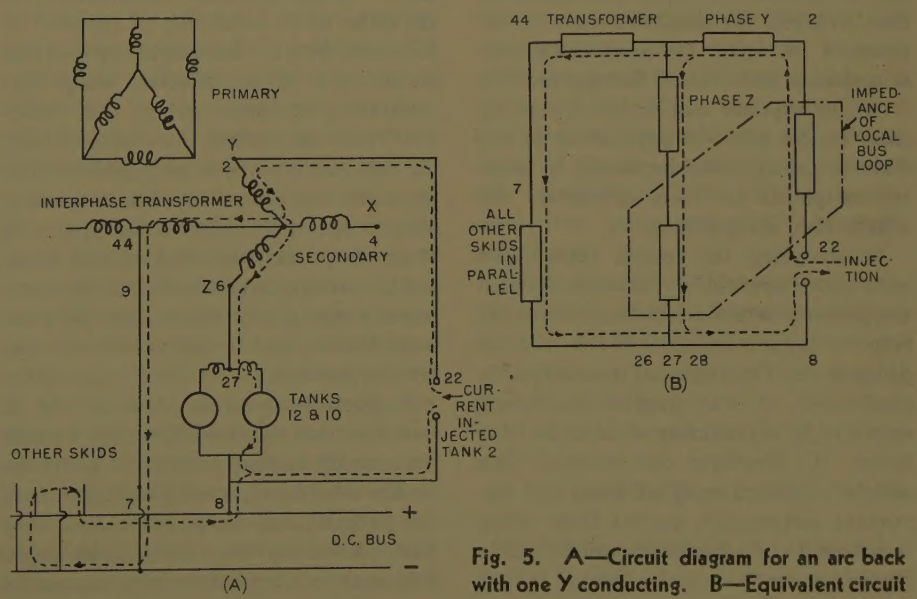


Fig. 5. A—Circuit diagram for an arc back with one Y conducting. B—Equivalent circuit

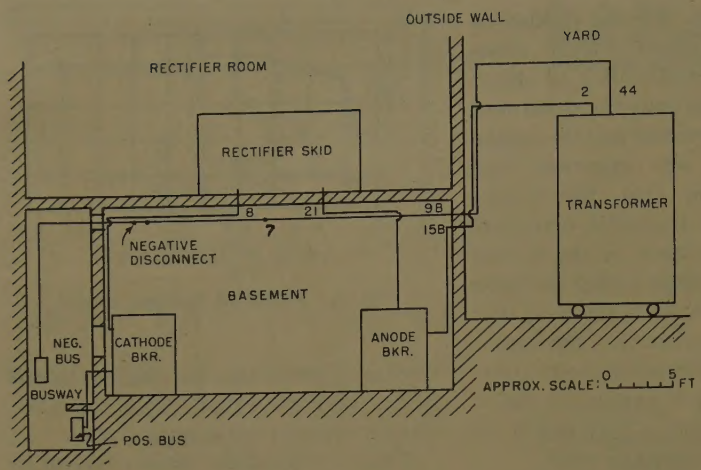


Fig. 6. Physical layout of power system

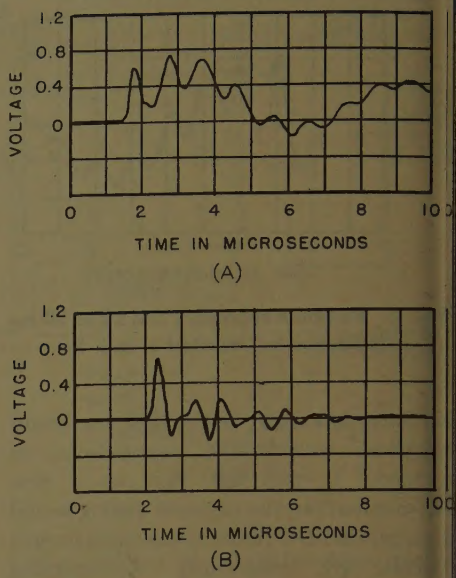


Fig. 7. Response to 20 ma injected at tank 2

A—At points 9B and 15B
B—At points 7 and 21 with short circuit between points 9B and 15B

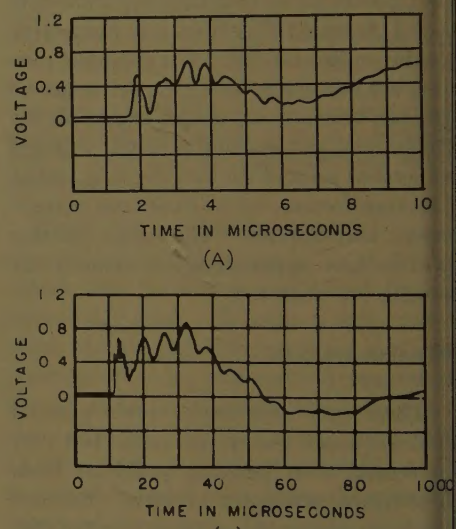


Fig. 8. Response at points 2 and 44 to 20 ma injected at tank 2

scope sweep somewhat ahead of the pulse but in some of the applications where greater physical distances were involved, it was found more convenient to reverse the process, i.e., the pulse generator was triggered by a signal emanating from the sweep generator of the oscilloscope.

A minor irritation was occasioned by interference with the electron beam in the oscilloscope caused by the strong d-c magnetic fields set up by the current in a neighboring d-c bus. This could be reduced somewhat by aligning the beam axis as closely as possible with the direction of the magnetic field. It was ultimately reduced to negligible proportions by putting the oscilloscope in a welded steel box made from a 1/4-inch boiler plate.

Analysis of Results

As explained in the section "Current Arrangements and Test Procedures," the procedure in test was to connect jumpers across all tanks assumed to be conducting at the instant under consideration, then to inject the simulated chop at the point of interest. Different circuit arrangements were tested in turn and for each the response was explored, starting close to the point of application of the disturbance and working away to more remote points to where ultimately the effects were insignificant.

In analyzing the results below, this same procedure will be followed. Which components were contributing most to the response became apparent as the program progressed. Once this had been definitely established, it was possible to reduce appreciably the number of tests in later series. To illustrate, one series of tests will be discussed in detail here, and important information gained from other series will be discussed briefly. The particular condition to be treated is that in which a surge follows an arc-back with one Y only conducting. No special significance attaches to this condition, but it serves well to illustrate the behavior of the system, and demonstrates the method.

It is assumed that the current suppression occurs in the rectifier that has arced back. A rough circuit diagram of the condition is shown in Fig. 5(A), where the manner in which the injected current divides is indicated. Fig. 5(B) is an approximate equivalent circuit. Points in the circuit are identified by the same numbers as those used in Fig. 1; the physical location of these points can be understood by the layout as shown in Fig. 6.

Consider first the trace observed between points 7 and 22, shown in Fig. 3. Since 7 and 22 are fairly close electrically to the points of injection, this oscillogram represents the response of most of the system. We notice several features: in the first place it is multifrequency. Second, it reaches its crest in approximately 0.3 μsec; therefore one of its important component frequencies is quite high. Third, the peak magnitude is of the order of 60 volts/amp of chopped or injected current. One might say that this corresponds to an effective surge impedance of 60 ohms; we call this "effective" because, as has been pointed out, the response is really a composite one, the summation of several separate responses; also, it inevitably includes damping. In the next few oscillograms, it will be possible to investigate as to what circuit components are giving rise to the different constituents, and to estimate their relative importance.

A quick check across points 21 and 22 indicated that the dividing reactor located between the tanks contributed very little. On the other hand, it was found that there was a remarkable change at location 9*B* to 15*B*, shown in Fig. 7(A). The high-frequency component is quite drastically reduced but the lower frequency remains virtually unchanged. This is very informative; it means that the high frequency is generated, for the most part by the comparatively short lengths of bus between the rectifier skid and the basement wall and the negative disconnect (the *B* in the designations 9*B* and 15*B* indicates that the measuring points are in the basement; see Fig. 6).

We conclude that the lower frequency is the product of equipment beyond the basement wall, namely, the interphase transformer and the power transformer outside in the yard. This was further confirmed and pinpointed by subsequent

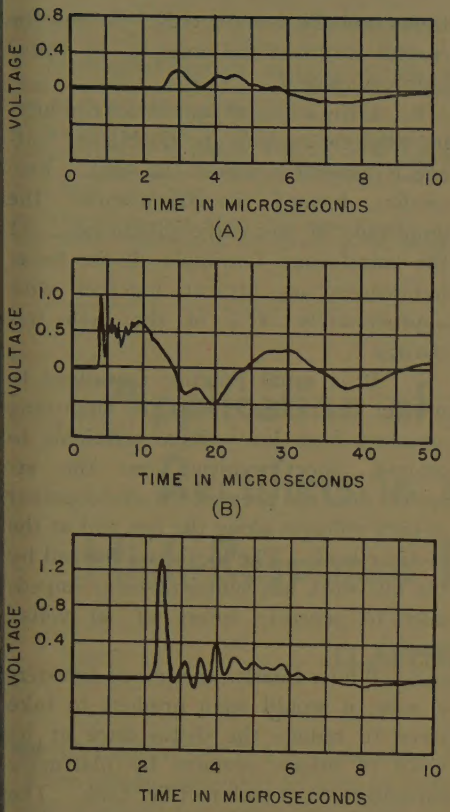


Fig. 9. Response to 20 ma injected at tank 2

A—At points 7 and 8
B—At points 7 and 21 with 13.8-kv winding of transformer open-circuited
C—At points 7 and 8 injected at these two points

tests. Points 9*B* and 15*B* were short-circuited, and the response observed across points 7 and 21; the effect is shown in Fig. 7(B). The low frequency has disappeared but the high frequency remains.

In a series of later tests current pulses were injected at the transformer terminals (points 2 and 44) and the response was measured at the same points; see Figs. 8(A) and 8(B) with different time scales. Now the lower-frequency component is evident, together with some high frequency; this arises from the length of bus from the transformer to the anode breaker, which could not be readily disconnected. Fig. 8(B) reveals a condition which could not be appreciated from the preceding oscillograms because of their time scales: the presence of a yet lower frequency component of significant proportions. It was thought that this was contributed by the interphase transformer, an assumption confirmed by obtaining the transient across two phases (points 2 and 6). This component was found to have disappeared.

A very significant fact was revealed by taking measurements at points 7 and 8. The voltage generated across the other 13 skids of the potline was thereby observable. This voltage is shown in Fig. 9(A) and is seen to be quite small relative to the voltage on the skid and transformer where the surge originates. This is not surprising because the other 13 skids are connected in parallel through the d-c bus and must therefore present a low surge impedance. It means that when a current suppression occurs, its effects are localized to the transformer and rectifier equipment immediately involved.

To recapitulate the results thus far, it may be said that a surge in a rectifier under the conditions prescribed excites three principal components. The highest in frequency, approximately 1 mc (megacycle), and amplitude, is caused by the bus run from the rectifier skid to the transformer and back to the negative bus. A quarter wavelength line with a natural frequency of 1 mc would be approximately 250 feet long, which would appear to be more extensive than our circuit. However, because of the shape of the circuit—an irregular, tortuous loop—such a comparison is probably a poor approximation. The transformer itself oscillates at around 165 kc; its amplitude is about 50% of the high-frequency constituent. A yet lower frequency of about 12.5 kc is generated by the interphase transformer; this component also has about 50% amplitude, but because of its low frequency it persists much longer, and for this reason it may be potentially more dangerous. Very little of the voltage transient reaches other skids on the potline; the effect of the disturbance is felt only close to its source.

Many other circuit conditions were investigated during the field tests; results followed very much the pattern of those described. For example, tests were made to simulate the effects of surge occurring at the time of an arc-back with both Y's conducting. The only difference between this circuit arrangement and that just described was that a phase in the second Y was also conducting.

The connection through the other Y was in parallel with all the other skids across the d-c bus. As this is already a comparatively low surge impedance, the extra parallel path could hardly be expected to make any dramatic difference to the over-all picture; this indeed proved to be the case. It was interesting to observe slight differences in the responses according to which phase in the second Y was carrying current. These differences arise because of the coupling, or lack of it (coupling is effected through the common primary winding), between the conducting phase and the arc-back phase.

Two other test series are worthy of comment. The first showed the consequences of a surge after the primary a-c line had been opened, the second recorded events consequent upon a current surge entering a rectifier skid from its unit bus. The first of these conditions could arise when a potline is being de-energized; it would prevail if a surge occurred during the interval between tripping out the a-c circuit breakers and opening the cathode breakers and negative disconnects.

Under these circumstances we would expect a difference in the low-frequency components of the response because the transformer presents a different surge impedance on open circuit than when short-circuited through the incoming supply.[5,6] In general, the surge impedance is higher and the natural frequency lower for the open-circuit condition; this was reflected in the oscillograms. Fig. 9(B) is typical, showing that the natural frequency of the open-circuited transformer is approximately 50 kc instead of the 165 kc observed in Fig. 8, with the primary winding short-circuited. The high-frequency oscillations remain sensibly the same, so that the first peak is comparatively unaffected. In general, we may say that the effect of opening the a-c line is to cause the higher-voltage peaks to persist longer through the constructive interference of the lower frequencies.

For studying the behavior of a rectifier skid when a surge enters from its unit bus, the skid concerned was isolated by opening the cathode breaker and the negative disconnect. The stimulus was applied and the response recorded at the unit bus, points 7 and 8. Two conditions were studied, normal operation and commutating. The transient excited when the circuit is in normal operation is shown in Fig. 9(C); the commutating condition was only slightly different because of the two parallel paths through the transformer. The oscillograms indicate the generation of 67 volts/amp suppressed.

Let us assume that there are three possible ways in which a current surge can enter from the unit bus:

1. It could be generated by a chop at the cathode breaker.
2. It could come from a neighboring skid.
3. It could be generated by the potline.

The three possibilities will be considered in turn.

To generate an overvoltage of 1 kv the circuit breaker must chop 15 amp; to generate 10 kv, 150 amp, etc. Ample oscillograms of the behavior of the cathode breakers have been taken and there is absolutely no evidence that they chop in this manner.

Regarding surges from neighboring skids, it may be said that all the skids in parallel would experience the surge; consequently the surge current would be only a small fraction of that in the surging skid. This means that the overvoltage appearing in the surging skid would be much greater than that in all of the other skids. Thus, we need not look for trouble from this source. Finally, a surge of current from the potline would divide at the bus into parallel paths through all the connected skids in a more or less equitable manner. It would therefore need a very considerable surge from the source to cause significant overvoltages at the unit bus and on the skids.

In full operation only 1/14 of the surge or suppression would be felt in any one skid or at any one transformer terminal. From Fig. 8(B) which shows the transformer's response to a chop, note that the peak voltage/amp suppressed is given by

$$Z_{0\,\text{effective}} = \frac{0.8}{0.02} = 40 \text{ volts/amp, approx}$$

To generate 35 kv, a figure that has actually been recorded, assuming the system is linear, would call for a suppression of

$$\frac{35{,}000}{40} = 875 \text{ amp}$$

But this is the current per skid. If the suppression emanates from the potline, it must have a magnitude of $14 \times 875 = 12{,}200$ amp in the line itself.

The very-high-frequency component of Fig. 7(B) does not contribute to the peak, being attenuated before that point is reached. However, the intermediate frequency is important, and to stimulate it significantly the current suppression should occur in a time less than half its period, i.e., in 2 or 3 μsec. It should be pointed out that because the low-frequency term makes the greatest contribution to the peak voltage at the transformer, a current suppression of the above magnitude taking as long as, say, 30 μsec could produce a surge voltage of 28 kv.

Conclusions

The authors believe that the method described is a very useful one for studying the behavior of complex industrial power networks when they are disturbed by surges of current. Repetitive current surging already applied successfully to individual components[8] is extended to an entire network.

The following, more detailed, conclusions apply specifically to the installation tested. However, in a general way, they are applicable to other similar potline systems. The method, of course, could be equally well applied to many other kinds of industrial power networks.

1. The overvoltages could be generated on the system by abrupt changes in current in any branch in any skid. The effect would be felt only locally, i.e., in the immediate Y to which the surging circuit was connected. Effects on the adjacent Y would be relatively slight, and on other skids, insignificant.

2. The effective surge impedances are in the range of 30 to 65 ohms; as an upper limit these could give rise to 30 to 65 volts/amp suppressed, so that to generate 35 kv a chop of at least 560 amp would have to occur.

As stated earlier, to determine the precise cause of surges was not the prime purpose of this investigation. Nevertheless, deductions from the results strongly suggest the rectifiers as the probably source. We say this because of the absence of evidence that the anode or cathode breakers chop currents of the order of several hundred amp, and also because the possibility of dangerous surges originating in the potline is all but discounted by virtue of the very low surge impedance presented to the line by the many parallel skids and transformers.

4. A surge in a rectifier would generate several frequencies of voltage in local circuits. The very high frequency (1 mc) would only appear as a significant constituent when the change of current was extremely fast, 0.3 μsec or faster. A somewhat lower voltage could appear across the transformer by current changes in 2 μsec or less. When conduction exists in only one Y of a skid, the interphase transformer can present a comparatively high surge impedance. This could give rise to voltage transients of the order of 35 volts/amp suppressed, and since the natural frequency of this component is low, the decay time of the current could be as long as 20 μsec. Saturation of the interphase transformer will tend to limit this voltage.

5. A surge in one rectifier tank with the adjacent tank (to which it is connected through the anode dividing reactor) conducting, would develop transient voltages in the local circuit comprising the two tanks and the reactor only. Across the reactor surges of the order of 20 volts/amp can appear.

6. If the adjacent tank is not conducting when the tank surges, the highest voltage is generated across the tank. Proceeding toward the transformer, the amplitude of the surge diminishes. At the transformer terminals all the lower frequencies are present together with approximately 50% of the high frequency.

7. It is usual practice nowadays to protect the transformers by lightning arresters from low-voltage terminals to ground. Short-circuiting out the arresters does not prevent the development of high voltages along the bus and at the rectifier tank. The local loop formed by the bus work has sufficient surge impedance to generate spikes of 50 volts/amp.

8. While, clearly, the use of arresters is wise, it would seem prudent to take steps to reduce the disturbance at its point of origin, perhaps by placing a capacitor across each rectifier tank. The surge impedance of the bus work is close to 100 ohms (estimated from the oscillograms making due allowance for the apparent damping) and the natural frequency is of the order of 1 mc. These figures correspond to an equivalent L and C of 1.6 μhenrys and 0.0016 μf. A protective capacitor C of about 2.5 μf would seem appropriate.

This value, which may seem high, is suggested for the following reason: A surge of 35 kv is reported to have been measured before lightning arresters were installed. The points of measurement were the anode breaker and the negative bus (points 9B and 15B) quite close to where the arresters were subsequently located. If this was due to a current surge at a rectifier, we may expect that the surge voltage at the skid would be perhaps 50% greater, say, 50 kv. Augmenting the stray capacitance by 2.5 μf will reduce the surge impedance by a factor of 40, so that the surge peak will be reduced to 1.25 kv.

If a higher level than this is acceptable then, of course, less capacitance could be used. C' should have in series with it a resistance just sufficient to keep within reasonable bounds the discharge current from the capacitor when the rectifier fires. Otherwise the connections to C' should be kept as short as possible.

References

1. Load-Dropping Tests on a Large Ignitron Rectifier Installation, S. L. Pope, J. H. Dillard, C. R. Marcum. *AIEE Transactions*, pt. I (Com

munication and Electronics), vol. 72, May 1953, pp. 164–75.

2. Arc-Backs in Rectifier Circuits—Artificial Arc-Back Tests, R. D. Evans, A. J. Maslin. *Ibid.*, vol. 64, June 1945, pp. 303–11.

3. Field Tests on a 100-Megawatt Rectifier Installation, J. K. Dillard, J. Keifer, C. S. Hague. *Ibid.*, pt. I (*Communication and Electronics*), vol. 74, May 1955, pp. 245–53.

4. The Cause of High Voltage Surges in Rectifier Circuits, A. W. Hull, F. R. Elder. *Journal of Applied Physics*, New York, N. Y., vol. 13, June 1943, pp. 372–77.

5. Voltage Transients Due to Arc Extinction, H. C. Steiner, R. W. Strecker. *AIEE Transactions*, pt. I (*Communication and Electronics*), vol. 79, May 1960, pp. 139–44.

6. The Effect of Current Chopping in Circuit Breakers on Networks and Transformers. I—Theoretical Considerations, T. H. Lee. *Ibid.*, pt. III (*Power Apparatus and Systems*), vol. 79, Aug. 1960, pp. 535–44.

7. *Discussion* by K. J. R. Wilkinson of The Recurrent-Surge Oscillograph and Its Application to the Study of Surge Phenomena on Transformers. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 96, pt. II, no. 50, Apr. 1949, pp. 269–81.

8. The Effect of Current Chopping in Circuit Breakers on Networks and Transformers. II—Experimental Techniques and Investigations, Allan Greenwood. *AIEE Transactions*, pt. III (*Power Apparatus and Systems*), vol. 79, Aug. 1960, pp. 545–55.

Application of Silicon-Controlled Rectifiers in a Transistorized High-Response D-C Servo System

CLARENCE CANTOR
NONMEMBER AIEE

THE COMMERCIAL introduction of the silicon-controlled rectifier in early 1958 disclosed the possibilities of its use in many applications usually performed by thyratrons, magnetic amplifiers, and rotating amplifiers.[1,2] One such application, which so far has not been thoroughly explored, is the use of these rectifiers as the power amplifier in a d-c servo system. This paper presents the results of designing, building, and testing a servo system which incorporates the controlled rectifiers as the power amplifier for the d-c servomotor. The final aim is to evaluate the performance of the controlled rectifier in this application.

The silicon-controlled rectifier is basically a p-n–p-n semiconductor with three rectifying junctions, in which regenerative switching is achieved through an avalanche breakdown of the center p-n junction. Breakdown occurs either by increasing the anode voltage to the breakover point or by introducing a low-level gate to cathode current with the anode positive but at less than the breakover voltage. Naturally, the latter method is used for power control applications. The rectifier regains its blocking state when the anode voltage drops down close to zero. This method of control is very similar to that used for thyratrons, except that the latter requires voltage control while the controlled rectifier uses a current control signal. In addition, because of junction heating effects and variation of firing current requirements with temperature, current pulses are more desirable than steady currents as a means of firing the controlled rectifier.

The controlled rectifier has the characteristics of small size, fast switching action, low voltage drop (about 1 volt), and high current rating, up to 16 amp (amperes).[3,4] The thyratron, while possessing fast response and high current capacity, is bulky and has a relatively high voltage drop (10 volts or more). Thus, there is considerable incentive in investigating the performance of the controlled rectifier in applications currently performed by thyratrons, including the one investigated in this paper. Except for the firing control circuitry, much of the literature on thyratron control of d-c motors[5,6] is applicable to the controlled rectifier, and offers a good starting point when the application of the controlled rectifier to a d-c servomotor is considered.

The basic circuit selected for this investigation is a half-wave bidirectional rectifier circuit. While producing less output current and more ripple than a full-wave circuit, the half-wave circuit is simpler, and adequate for d-c servomotors up to about 1/2 horsepower.

Paper 60-864, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Summer General Meeting, Atlantic City, N. J., June 19–24, 1960; and re-presented for discussion only at the Fall General Meeting, Chicago, Ill., October 9–14, 1960. Manuscript submitted November 4, 1959; made available for printing August 19, 1960.

Clarence Cantor is with the U. S. Naval Weapons Plant, Washington, D. C.

This paper is a summary of the author's Master of Science thesis (1959) at the Massachusetts Institute of Technology, where the facilities of the Electronic Systems Laboratory were utilized for the experimental work. Prof. George C. Newton was the thesis supervisor. The Office of Naval Research, under contract *NOnr-1841-(53)*, and the U. S. Naval Weapons Plant sponsored the research.

As mentioned previously, pulse firing is desirable for the silicon-controlled rectifier. The basic control scheme used in the present investigation is pulse firing from unijunction transistors, which have the capability of performing as a relaxation oscillator. This feature dovetails nicely into a simple pulse-firing scheme for the controlled rectifier.

The circuitry and controls selected for the present system do not by any means represent an optimum, but rather an attempt to achieve a fairly simple system giving high servo performance. This is a very effective means of uncovering limitations, if any, imposed on servo performance by the silicon-controlled rectifiers themselves.

System Operation

Fig. 1 is a schematic diagram of the over-all system. The servomotor is a small d-c torque motor rated at 115 volts, 0.58 amp, 0.23 lb-ft (pound-foot). The motor is fed from the silicon-controlled rectifier circuit, which uses half-wave bidirectional control. The supply voltage is 115 volts, single-phase, 400 cycles. Two control signals determine the output of the rectifier circuit. One is the 400-cycle a-c positional error signal from the control transformer. The other is a d-c stabilizing signal from the output of the tachometer amplifier and stability network. The latter signal provides velocity feedback through a high-pass filter. The source of this signal is a d-c tachometer generator directly coupled to the torque motor. A control trans-

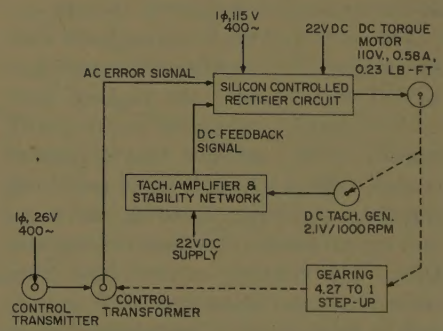


Fig. 1. Schematic diagram

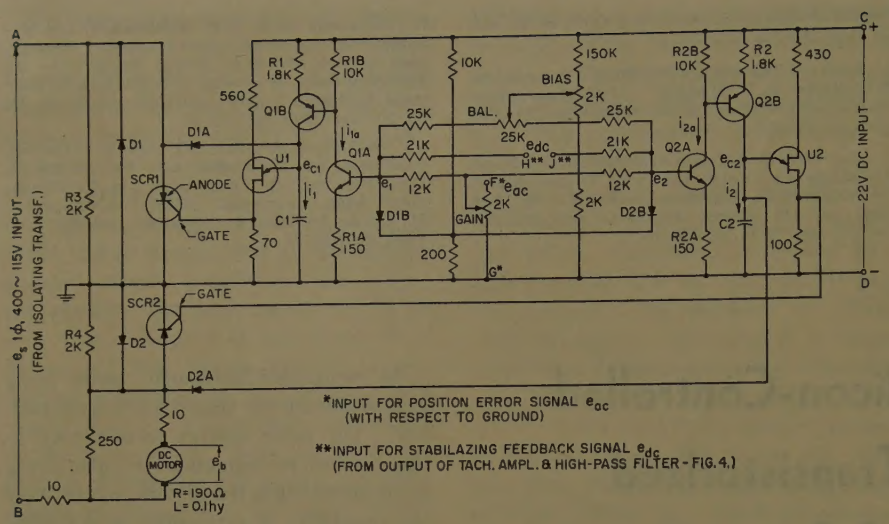


Fig. 2. Silicon-controlled rectifier circuit: power and control

mitter, identical to the control transformer, serves as the position reference for the servo system.

The heart of the system is the silicon-controlled rectifier circuit, a circuit diagram of which is shown in Fig. 2. The rectifiers and motor are connected in a half-wave bidirectional circuit supplied at A and B by 115 volts, single-phase, 400 cycles. Power diodes $D1$ and $D2$ are used to permit a common cathode connection of the silicon-controlled rectifiers. This permits some simplification of the control circuitry. A 22-volt d-c supply is fed across terminals C and D. This provides power for the transistor circuitry.

The firing of silicon-controlled rectifiers $SCR1$ and $SCR2$ is initiated by positive current pulses from unijunction transistors $U1$ and $U2$ respectively. (For unijunction transistor characteristics, see references 7 and 8.) The range of firing control for each rectifier is from 10 to 180 degrees. The relative firing angles of $SCR1$ and $SCR2$ determine the magnitude and direction of the average motor current and hence torque. For zero torque, both $SCR1$ and $SCR2$ are prevented from firing until late (160 degrees) in their respective half-cycles. Some conduction of both rectifiers is desirable at standstill in order to obtain greater sensitivity at null.

The control of the firing angle for each silicon-controlled rectifier is accomplished as follows: Consider the positive half-cycle of supply voltage (A positive with respect to B). Assume that a positive control voltage e_1 (made up of positional signal e_{ac}, stabilizing signal e_{dc}, and a d-c bias signal) exists at the base of transistor $Q1A$. The collector current from transistor $Q1A$, and hence from $Q1B$ as well, will be approximately proportional to e_1. Thus the rate of build-up of the voltage e_{c1} across condenser $C1$, during the positive half-cycle, is instantaneously proportional to the control voltage e_1. When the voltage across $C1$ builds up to 12 volts, $U1$ discharges $C1$ into the gate of $SCR1$, causing $SCR1$ to conduct.

When $SCR1$ fires, $C1$ is short-circuited by $SCR1$ via diode $D1A$. During the negative half-cycle, e_{c1} remains clamped at about zero volts via leakage resistor $R4$ and diodes $D1$ and $D1A$. When terminal A again becomes positive, the condenser voltage e_{c1} can build up again in accordance with the net control voltage e_1. The net result is that the firing angle of $SCR1$ is inversely proportional to the average value of control voltage e_1 up to the point of firing, provided e_1 is positive during the positive half-cycle. For e_1 negative, no build-up would occur. With control signals e_{ac} and e_{dc} equal to zero, e_1 has a d-c bias to provide a charging rate into $C1$ such as to fire $SCR1$ at about 160 degrees.

From the symmetry of the circuit it can be seen that whatever applies to control voltage e_1 during the positive half-cycle also applies to control voltage e_2 during the negative half-cycle. Thus the firing angle of $SCR2$ is inversely proportional to the average of e_2, up to the point of firing. Control voltage e_2 also has a d-c bias that is sufficient by itself to fire $SCR2$ at about 160 degrees.

Fig. 3 shows some circuit waveforms for various combinations of control signals e_{ac} and e_{dc}. For the condition of zero control signals, as shown in Fig. 3(A), both $SCR1$ and $SCR2$ fire at the same angle (160 degrees) so that the net torque developed is zero. Fig. 3(B) illustrates the case of a forward (in-phase) error signal e_{ac}, with e_{dc} and back electromotive force (emf) e_b equal to zero. During the positive half-cycle, when condenser $C1$ is unclamped, e_{ac} is positive and causes condenser voltage e_{c1} to build up at a faster rate than the zero-signal rate. This results in an advancement of the firing angle for $SCR1$. During the negative half-cycle, when condenser $C2$ is unclamped, e_{ac} is negative and causes the firing of $SCR2$ to be eliminated. Thus a forward torque is developed, tending to reduce the positional error signal e_{ac}. In Fig. 3(C), a reverse stabilizing signal e_{dc} is shown added. Since a negative e_{dc} tends to decrease control voltage e_1 while increasing e_2, it has an effect opposite to that of the

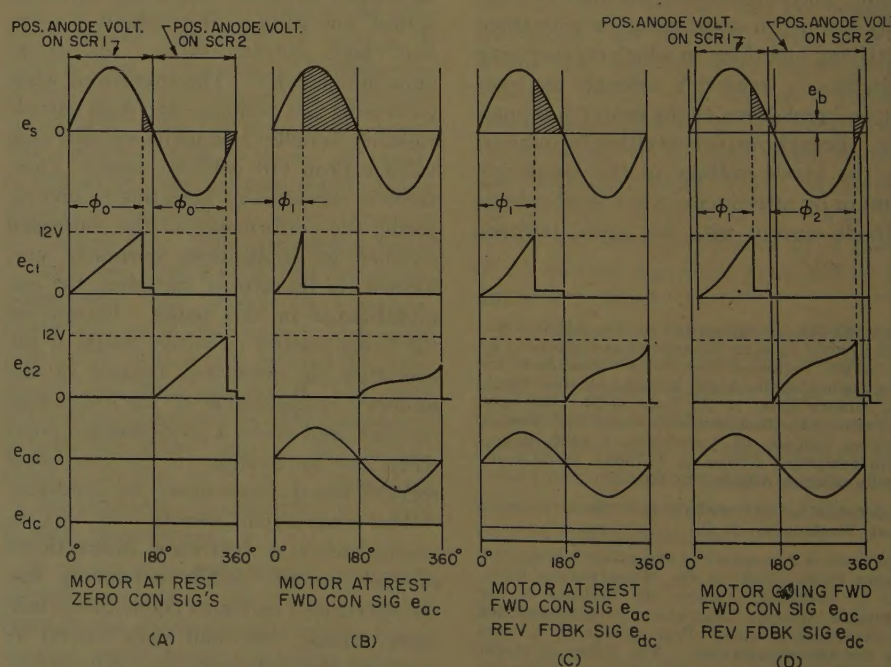


Fig. 3. Rectifier circuit waveforms

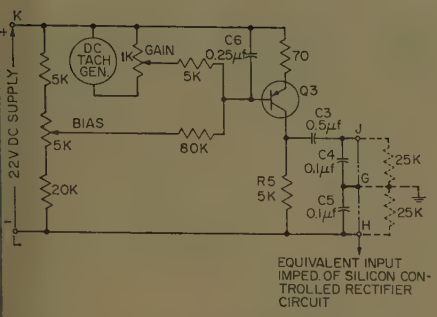


Fig. 4. Tachometer amplifier and stabilizing network

forward error signal e_{ac}. For the case shown, e_{ac} has the larger effect and advances the firing of $SCR1$, though not as much as before. Thus the forward torque is reduced.

Fig. 3(D) illustrates the effect of motor back emf. The same magnitude of control signals e_{ac} and e_{dc} as in Fig. 3(C) is assumed. The presence of a negative back emf (forward motor rotation), delays the beginning of positive anode voltage on $SCR1$ while advancing it for $SCR2$. Since the firing angles are keyed with respect to the beginning of positive anode voltage, the firing of $SCR1$ is delayed while that of $SCR2$ is advanced. In addition, the same back emf cuts off $SCR1$ earlier while allowing $SCR2$ to conduct longer. Lastly, the instantaneous net voltage available to drive current through the motor armature resistance is less during $SCR1$ conduction and more during $SCR2$ conduction. The result is that $SCR1$ conducts less average current, while $SCR2$ conducts more, thus reducing the forward torque.

The circuit diagram for the tachometer amplifier and stabilizing network is shown in Fig. 4. The bias potentiometer is adjusted to obtain about 10 volts across transistor output resistor $R5$ thus permitting a voltage swing of ± 10 volts. The maximum amplifier gain is about 32. Condenser $C3$ together with the equivalent input resistance of the rectifier control circuit, forms a high-pass network of the form $T_1/(1+Ts)$ where T is approximately 0.03 sec (second). Condensers $C4$, $C5$, and $C6$ serve to minimize 400-cycle noise pickup as well as 400-cycle feedback from the torque motor via the direct-coupled d-c tachometer generator.

Speed-Torque Characteristics

The locked rotor torque characteristics can be readily derived on the basis of negligible armature inductance and rectifier voltage drop, with the zero-signal conduction angle ϕ_0 as a parameter. The resultant characteristics show a similarity for both a-c and d-c control signals up to a torque output equal to one half of maximum. This is expected since the rectifier output depends on the average of the control signal over the firing angle. The torque characteristic for $\phi_0 = 160$ degrees is about the best as far as linearity and smoothness of control are concerned; this is the angle selected as the zero signal conduction angle. For $\phi_0 = 160$ degrees, the torque curve can be considered linear up to about 40% of T maximum with saturation beginning thereafter. Locked rotor current measurements confirm this.

The speed-torque characteristics were measured rather than derived, because of the added computing difficulty caused by the 250-ohm motor-shunting resistor. The resistor has the desirable effect of essentially reducing the circuit inductance. This in turn results in an increased locked rotor torque as well as a smoother transition between positive and negative torque. However, the regenerative braking effect of the shunting resistor during the non-conducting part of the cycle modifies the speed-torque characteristics and causes a reduction in maximum speed. This is not objectionable for the present system, since the torque motor is not mechanically designed for high speeds.

Measured speed-torque characteristics for low speeds and torque with $\phi_0 = 160$ degrees, are shown in Fig. 5. T_{max} corresponds to maximum locked rotor torque and for this system is equal to 0.07 pound-foot. The theoretical speed corresponding to a motor back emf equal to the peak of the supply voltage is defined as Ω maximum. For this system $\Omega_{max} = 280$. rad/sec (radians per sec). The a-c control signal α and d-c control signal β are normalized signals at the input to transistors $Q1A$ and $Q2A$. They are defined as

$$\alpha = \frac{2}{\pi} \times \frac{e_a}{e_0}$$

and

$$\beta = \frac{e_d}{e_0}$$

where e_a is the peak of the a-c signal component, e_d is the d-c signal component, and e_0 is the d-c bias signal required to produce ϕ_0, all with respect to ground at the input to transistors $Q1A$ and $Q2A$. This representation, together with the similarity of characteristics produced by small a-c and d-c control signals at low speed, permits the speed-torque curves to be shown as identical for both types of signals. The conversion of rectifier input signals e_{ac} and e_{dc} to α and β respectively can be done from the measured relationships:

e_{ac} (rms) $= 0.07\ \alpha$ volt

$e_{dc} = 0.3\ \beta$ volt

At the low speeds and torques shown in Fig. 5, the characteristics are essentially straight, parallel lines. Also for $-0.6 \leq \alpha$ (or β) ≤ 0.6, the lines are approximately equally spaced. Thus, for this range of control signals and for $-0.15 < \Omega/\Omega_{max} < 0.15$, the system may be considered linear.

Servo Analysis and Performance

An approximate frequency response for the rectifier circuit (locked rotor) can be derived under the assumptions that the signal frequency is small compared to the supply frequency, that only one of the

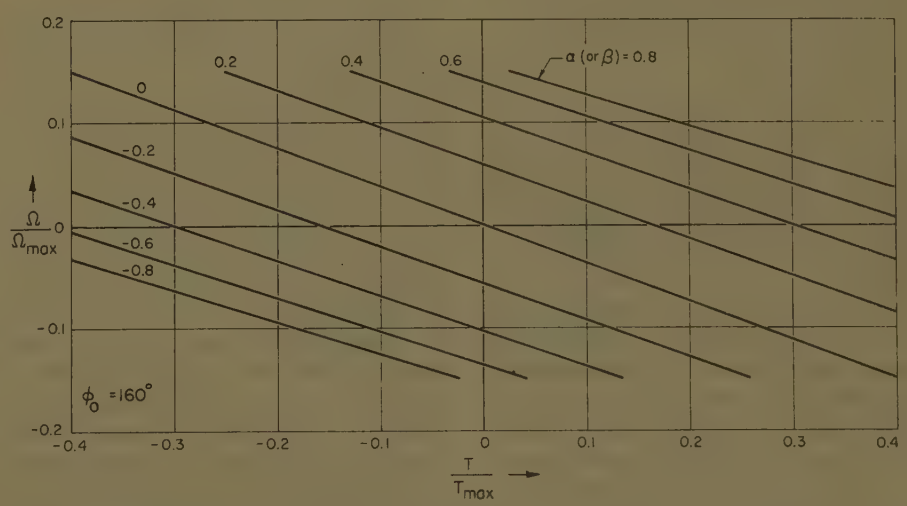


Fig. 5. Measured speed torque curves, low speeds, a-c control signal α, or d-c signal β, as a parameter

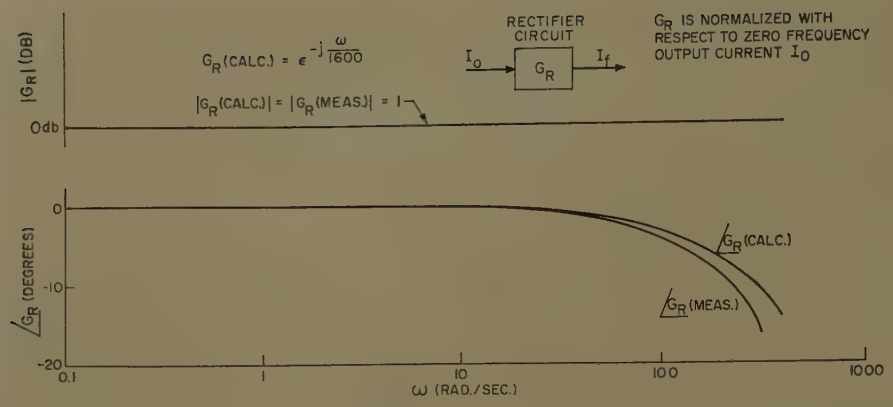


Fig. 6. Rectifier circuit frequency response

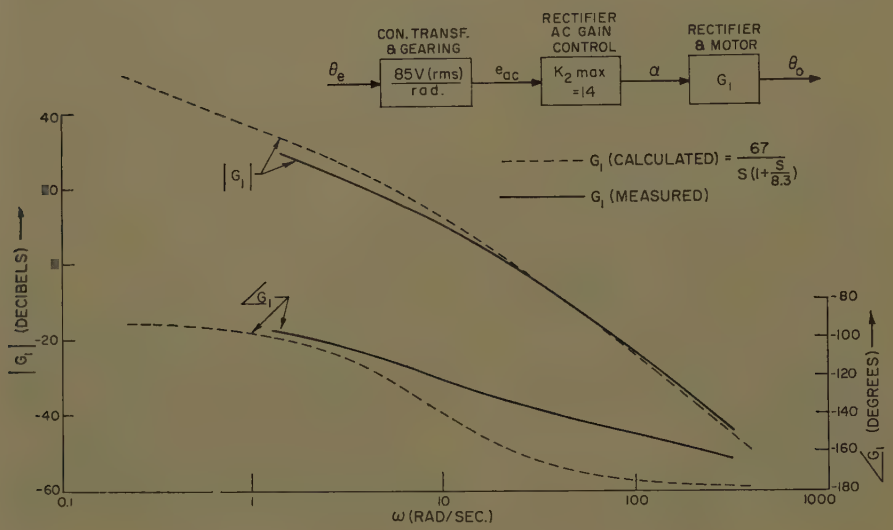


Fig. 7. Uncompensated open-loop frequency response

rectifiers is conducting during any one cycle of supply frequency, and that this conduction is small. Thus as an approximation, we may consider the conduction to be a current pulse at the end of the half-cycle over which the input control signal is averaged. The final result is

$$G_R=\frac{I_f}{I_0}=\epsilon^{-j\omega T/4}$$

where T is the period of the supply frequency, I_f is the fundamental component (signal frequency ω_c) of the rectifier circuit output current, and I_0 is the average output current at zero signal frequency (d-c). The same result is obtained for larger rectifier outputs that are still proportional to the input signals, if it is assumed that the conduction area can be approximately replaced by a current impulse of equal area located at the center of conduction.

The fundamental I_f at signal frequency ω_c is, strictly speaking, not a fundamental except when the supply frequency ω_s is an integral multiple of ω_c. However, even at in-between frequencies for ω_c, the predominant harmonic will be the one at signal frequency ω_c, as long as ω_c is small compared to ω_s. This major component is referred to as the fundamental I_f, to which $G_R=\epsilon^{-jT\omega/4}$ applies. The 1/4-cycle delay of this approximate frequency response is due to the averaging used in the control scheme and to the characteristics of the half-wave circuit. It should be noted that this delay does not apply to sudden inputs such as a step function, where the delay time can be as much as one cycle of the supply frequency.

The calculated frequency response of G_R is shown in Fig. 6. Also shown is the measured frequency response of the rectifier circuit. The amplitude response is the same for both, namely unity. The phase response of G_R (calculated) and G_R (measured) shows good correlation, considering the approximations used in deriving G_R (calculated). In any case the phase lag of the entire rectifier circuit is very small up to 50 cps (cycles per sec) which is one eighth of the supply frequency.

Assuming that the motor load is pure inertia, that the armature inductance is negligible, and that the lag of the rectifier circuit is negligible, the uncompensated system open-loop response for small signals and speeds can be obtained from the speed-torque curves of Fig. 5. This proves to be

$$G_1=\frac{\theta_0}{\alpha(\text{or }\beta)}=\frac{67}{s\left(1+\frac{s}{8.3}\right)}$$

The amplitude and phase response of this calculated G_1 are shown in Fig. 7; also shown is the measured frequency response. The differences in amplitude and phase are accounted for by the coulomb friction present in the system. The measured response is a further verification of the negligible phase lag of the rectifier circuit itself, at frequencies up to 50 cps.

An analytic approximation of the measured G_1 is

$$G_1=\frac{40}{s\left(1+\frac{s}{15}\right)}$$

An approximation of the compensated open-loop response G is

$$G\approx G_1 \text{ for } G_1H_1<1$$

$$G\approx\frac{1}{H_1} \text{ for } G_1H_1>1$$

In order that the compensated open-loop response may have a −20-db (decibels)/decade slope from about $\omega=33$ to $\omega=500$ rad/sec, the minor loop feedback H_1 is selected to be

$$H_1=\frac{0.025s^2}{\left(1+\frac{s}{33}\right)}$$

An asymptotic approximation of the compensated open-loop response G is shown in Fig. 8. It is desired that the complete open-loop response KG cross the zero db line at about $\omega=350$ rad/sec. This should result in a system of high bandwidth and good stability. Therefore K is adjusted to a value of 300, or $K_2=3.5$.

The measured closed-loop frequency response of the system for small signals is shown in Fig. 9. The response is flat to 43 cps, reaching a peak of 1.4 (3 db) at 55 cps, corresponding to the 90-degree lag point. The ±3-db bandwidth is about 65 cps. This response correlates fairly well with the previous open-loop measurements and calculations.

Typical transient response measurements are shown in Fig. 10. The 1/4-degree step function represents operation

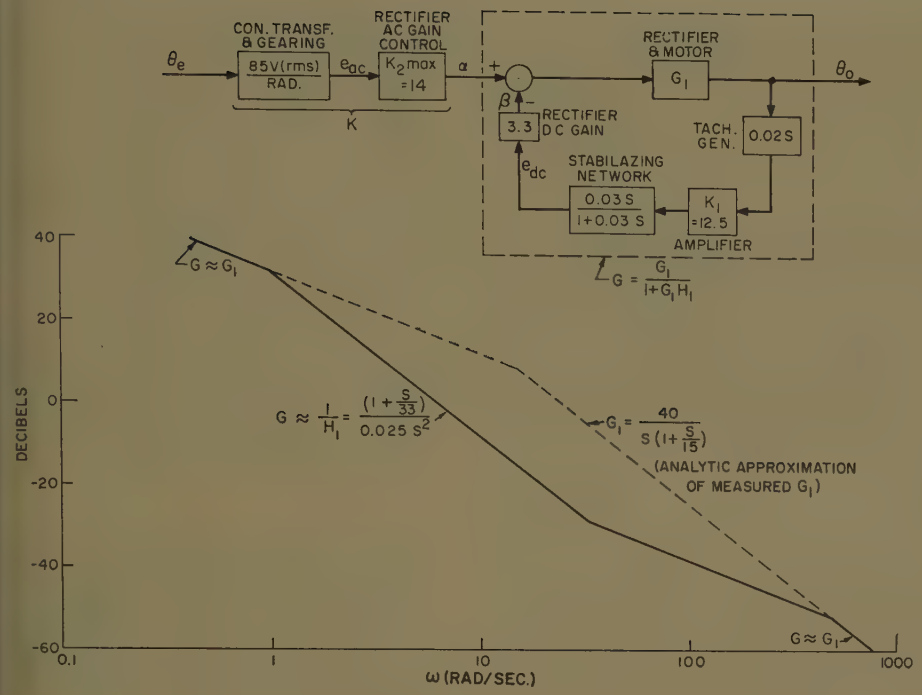


Fig. 8. Compensated open-loop frequency response, asymptotic approximation

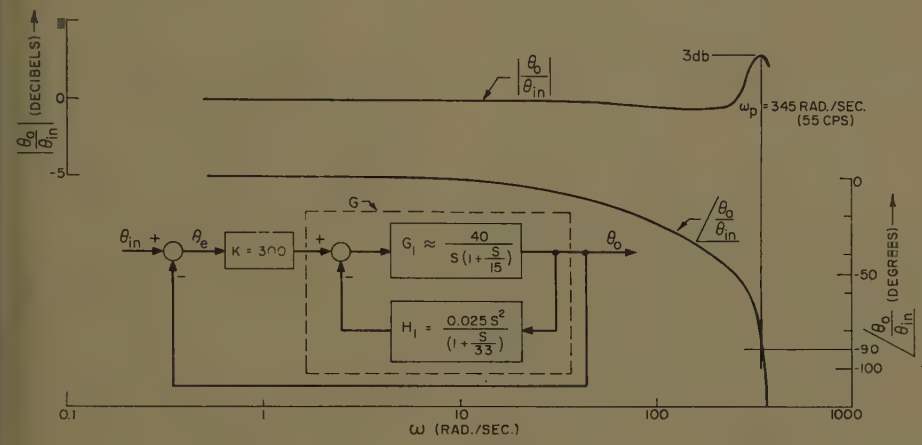


Fig. 9. Closed-loop frequency response, measured

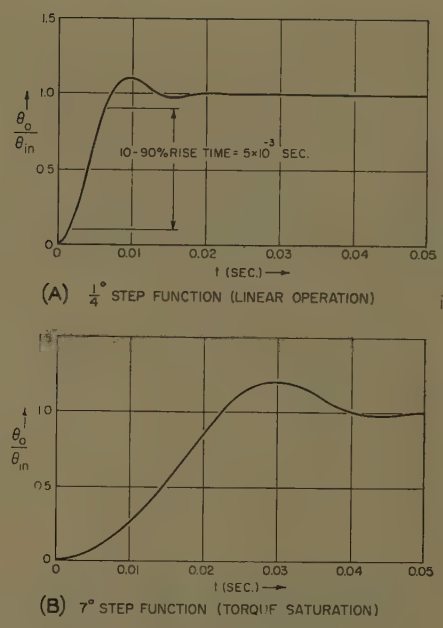


Fig. 10. System transient response

within the approximate linear range of the rectifier circuit. The 5-millisecond rise time is 0.33 of the period of the 65-cps bandwidth, which is reasonable. The 7-degree step function represents torque saturation through most of the response. This signal does not produce velocity saturation. A step function of 200 degrees or larger would be required to reach velocity saturation and this is not feasible with the single high-speed synchro in the existing system. However, other tests indicate that the maximum slewing velocity would be about 60 rad/sec. It should be noted that the transient response measurements do not include an initial delay time, which for the half-wave circuit can be as much as one cycle of the 400-cycle supply frequency.

The measured velocity constant of the system is 10,000 sec^{-1}. This checks closely with the velocity constant of 12,000 sec^{-1} obtained from the open-loop response of Fig. 8 (with $K=300$).

The torque constant of the system, with respect to the motor shaft, is 15-lb-ft/rad. The coulomb friction in the system is about 0.015 lb-ft. Thus the theoretical motor position error due to friction is 1×10^{-3} rad or about 0.06 degree. The actual position error is much less than this due to the 400-cycle dither inherent in the system. The measured friction error is about 0.01 degree.

Conclusions

The use of silicon-controlled rectifiers as the supply for a d-c servomotor permits the design and realization of a servo system of very high performance. As far as servo performance is concerned, there is no time lag associated with it other than that inherent in the control scheme and half-wave rectifier circuit. It is possible, using these rectifiers, to achieve a servo bandwidth up to one sixth of the supply frequency, and in all probability a bandwidth up to one fourth of the supply frequency is feasible.

The experimental system is designed around a small d-c torque motor. The same basic circuit, however, could be used for d-c servo motors up to about 1/2 horsepower, assuming that the motor itself had suitable servo characteristics. About the only changes required would be an increase in the heat sink capacity for the silicon-controlled rectifiers and an increase in the rating of the power diodes, such as to match the current requirements of the motor. The control circuit could remain very much the same.

The pulse-firing control scheme used gives very fast, reliable performance. During many hours of operation, there was no evidence of misfiring. The breakover voltage of the unijunction transistors remains stable at 12 volts when used with a constant 22-volt d-c supply. In the case of both the d-c and carrier-modulated control signals, the variable rate-charging circuit used gives a sharp intersection of the condenser voltage wave with the breakover voltage, thus permitting precise firing. The flexibility afforded by the use of either d-c or carrier-modulated control signals is another advantage of the circuit.

The main disadvantage of the firing control scheme is the measured 30-degree variation in zero signal firing angle during

circuit warm-up, due to variations in germanium transistor leakage with temperature. This produces a change in rectifier circuit gain, but very little drift since the balanced circuitry minimizes rectifier output drift. The variation in zero-signal firing angle can be minimized by the use of a lower transistor amplifier gain or by the use of silicon transistors.

The d-c servo performance obtainable with silicon-controlled rectifiers is at least equal to that obtainable with thyratrons. This together with the much smaller size and lower voltage drop of the controlled rectifier, makes it superior to the thyratron for use as the power amplifier in a d-c servo system.

References

1. Solid-State Thyratron Switches Kilowatts, R. P. Frenzel, F. W. Gutzwiller. *Electronics*, New York, N. Y., Mar. 28, 1958.

2. The Controlled Rectifier: Key to the Continuing Control Renaissance, J. D. Harnden, Jr. *AIEE Transactions*, pt. I (*Communication and Electronics*), vol. 77, 1958 (Jan. 1959 section), pp. 1006–19.

3. A Silicon-Controled Rectifier, I—Characteristics and Ratings, D. K. Bisson, R. F. Dyer. *Ibid.*, vol. 78, May 1959, pp. 102–06.

4. The Silicon Controlled Rectifier, F. W. Gutzwiller. *Electrical Manufacturing*, New York, N. Y., 1958.

5. Control System Components (book), J. E. Gibson, F. B. Tutuer. McGraw-Hill Book Company, Inc., New York, N. Y., 1958.

6. Control Engineers Handbook (book), J. G. Truxal. McGraw-Hill Book Company, Inc., 1958.

7. More Circuits Transistorized with the Silicon Unijunction Transistor, S. R. Brown, T. P. Sylvan. *Electronic Design*, New York, N. Y., pt. I, Jan. 8, 1958; pt. II, Jan. 22, 1958.

8. General Electric Transistor Manual, General Electric Company, Schenectady, N. Y., third edition, 1958.

9. PNPN Transistor Switches, J. L. Moll, M. Tannenbaum, J. M. Goldey, N. Holonyak. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 44, Sept. 1956, p. 1174.

Variable 3-Phase Reversible Voltage Sources With Use of Saturable Reactors

M. BOLT NONMEMBER AIEE A. SIMEON NONMEMBER AIEE W. SHEPHERD ASSOCIATE MEMBER AIEE

Synopsis: This paper describes the circuit arrangements, with appropriate analyses, by which two or three saturable reactors may be used to supply a 3-phase load. The output voltages of the devices are constant-frequency 3-phase, almost sinusoidal sources which may be smoothly and continuously varied from an almost balanced positive sequence to an almost balanced negative sequence by variation of the saturable-reactor control currents.

SEVERAL recent publications[1–5] describe investigations in which the polyphase wound-rotor induction motor has been used as the basis of stepless a-c variable-speed drives. Each of these schemes is characterized by the use of some form of static magnetic power modulator which supplies an induction motor. References 1 and 2 describe schemes in which saturable reactors are used in series with individual machine windings, while the scheme of reference 3 involves the use of static magnetic frequency doublers. L. R. Foote, in reference 4, has demonstrated the novel idea of using saturable reactors as interphase magnetic couplings in order to produce certain desirable voltage unbalances for speed control purposes.

This paper demonstrates some extended possibilities by which the interphase coupling of saturable reactors can produce variable, controllable, 3-phase voltages suitable for the control of polyphase induction motors in situations which require a simultaneous application of positive- and negative-sequence power.

Paper 61-37, recommended by the AIEE Industrial Control Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 29–February 3, 1961. Manuscript submitted October 25, 1960; made available for printing November 18, 1960.

M. Bolt, A. Simeon, and W. Shepherd are with the University of Manitoba, Winnipeg, Man., Canada.

The authors are indebted to the National Research Council of Canada for financial support of the research program on which this paper is based.

The Basic Saturable Reactor Connection

Fig. 1 shows the basic saturable reactor interphase connection, classified for convenience as connection 1. Two power windings, having the turns ratio p to 1 are connected in series with phases B and C respectively, with the magnetic polarity as shown by the dots.

To facilitate the analysis the reactor is assumed to be ideal, having no losses and no leakage flux. When this ideal magnetic coupling has unity turns ratio, the self-inductances of each power winding and the mutual inductance between them are equal. The mutual reactance is symbolized by X in Fig. 1 and is a function of the reactor control current.

It is further assumed that the reactor core has an idealized magnetization characteristic of the type shown in Fig. 2. X is then variable from zero at saturation to infinity in the unsaturated region and has different finite values corresponding to every point on the curved transitional region of the magnetization characteristic.

These assumptions permit the circuit of Fig. 1 to be analyzed by conventional steady-state sinusoidal circuit analysis techniques. The validity of the assumptions is discussed later.

An outline of the analysis of connection

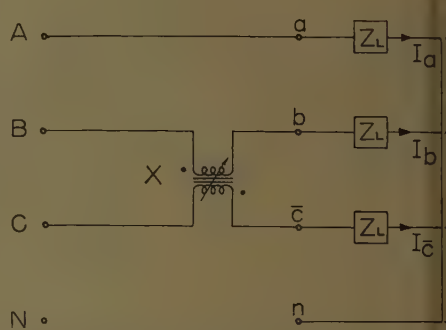


Fig. 1. Basic saturable reactor interphase connection 1

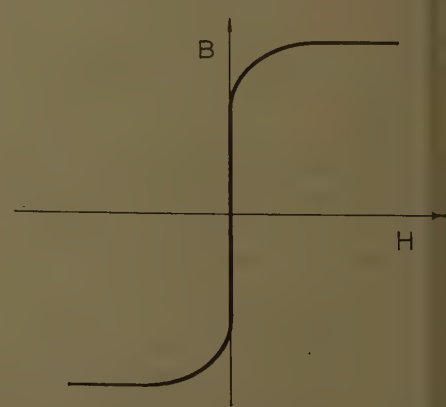


Fig. 2. Idealized magnetization characteristic of saturable reactor cores

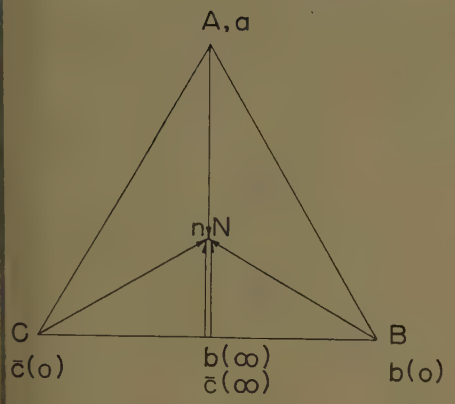


Fig. 3. Limiting load voltage vectors for basic connection 1

1 is given in Appendix I for the case of a balanced load, $Z_L = R_L + jX_L$ ohms per phase, and a unity turns ratio reactor.

Equations 10 and 11 from Appendix I are reproduced here for convenience; let

$$\tan \alpha = \frac{R_L}{2X + X_L} \tag{10}$$

$$V_{\bar{C}\bar{C}} = \frac{V_{CB}}{2}\left[\cos \alpha - \frac{X_L}{R_I} \sin \alpha\right] \underline{|\alpha} \tag{11}$$

The voltage vector diagrams for connection 1 are shown in Figs. 3 and 4. With zero direct-control current the reactor core has infinite permeability and X tends to infinity. The corresponding load voltages are denoted by a, $b(\infty)$, and $\bar{c}(\infty)$ in Fig. 3. With maximum direct-control current the reactor becomes saturated, having zero permeability and reducing X to zero. Load voltages V_{an}, V_{bn}, $V_{\bar{c}n}$, are then coincident with the applied voltages as denoted by the symbols a, $b(0)$, $\bar{c}(0)$, in Fig. 3. These two extreme voltage conditions are independent of the load impedance.

Taking point C of the voltage vector triangle ABC as origin, and line CB as the x-axis, equation 11 may be expressed in terms of the Cartesian co-ordinates x and y of voltage $V_{C\bar{c}}$.

$$\left(x - \frac{V_{CB}}{4}\right)^2 \left(y + \frac{V_{CB} X_L}{4R_L}\right)^2 = \left(\frac{V_{CB}}{4}\right)^2 + \left(\frac{V_{CB} X_L}{4R_L}\right)^2 \tag{12}$$

Equation 12 describes a set of circles having their centers on the perpendicular through point $V_{CB}/4$ in triangle ABC. The perpendicular displacement y_0 of the circle center from line CB and the circle radius r_0 are both functions of the load impedance.

With a purely resistive load, the loci of phase voltages V_{bn}, $V_{\bar{c}n}$ as X varies from zero to infinity are semicircles on the diameters $V_{CB}/2$ $V_{BC}/2$ respectively, Fig. 4(A) shows these loci with the phase voltage vectors at an arbitrary intermediate position of saturable reactor control current. Since, by equation 1, the load voltages must sum to zero and, by equation 6, V_{an} is constant, vectors V_{bn} and $V_{\bar{c}n}$ must lie along adjacent sides of a parallelogram with diagonal $N\bar{A}$.

When the balanced load is partially reactive, the voltage loci V_{bn} and $V_{\bar{c}n}$ will follow the circular paths described by equation 12 and illustrated in Figs. 4(B) and (C). The limiting condition of an entirely reactive load produces loci along the line of vector voltage V_{CB}. It is significant to note that the locations of the voltage loci described by equation 12 are independent of the saturable reactor reactance. For a fixed load impedance, the value of the saturable reactor reactance determines only the point of operation on a fixed locus.

Let a different constraint now be applied in such a way that the magnitude only of the load impedance is kept constant, instead of both the magnitude and the phase angle.

$$|Z_L| = \sqrt{R_L^2 + X_L^2} = \text{constant} \tag{13}$$

Combining equations 12 and 13 for a particular saturable reactor reactance X gives

$$\left(x - \frac{2X^2 V_{CB}}{4X^2 - |Z_L|^2}\right)^2 + y^2 = \left(\frac{V_{CB} X |Z_L|}{4X^2 - |Z_L|^2}\right)^2 \tag{15}$$

The loci of constant-load amplitude described by equation 15 are a set of eccentric circles centered along the line of voltage V_{CB}. Since the radii and the locations of the circle centers are functions of both the load impedance amplitude and of saturable reactor reactance X, it follows that, by controlling the value of X, a prescribed voltage unbalance can be sustained in the presence of load variation.

The voltage vector loci of Figs. 3 and 4 show that a load voltage phase inversion cannot be achieved by connection 1 for any R-L (resistance-inductance) load combination. The use of nonunity turns ratios with this connection results in different voltage loci but these show no advantage over the unity turns ratio performance described above.

The principle and analytical technique used for connection 1 is now extended to more complicated saturable reactor connections.

Reactor Connections Giving Reversible Output Voltages

One of the axiomatic requirements of a variable-speed drive is that the driving

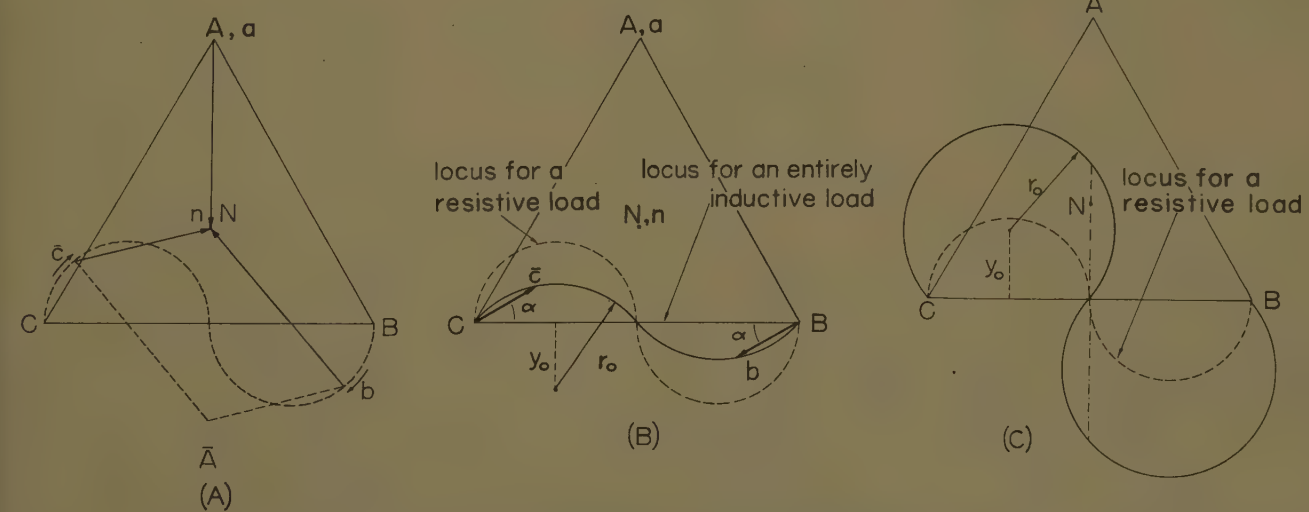


Fig. 4. Load voltage loci for basic connection 1

A—Resistive load B—Resistance-inductance load C—Resistance-capacitance load

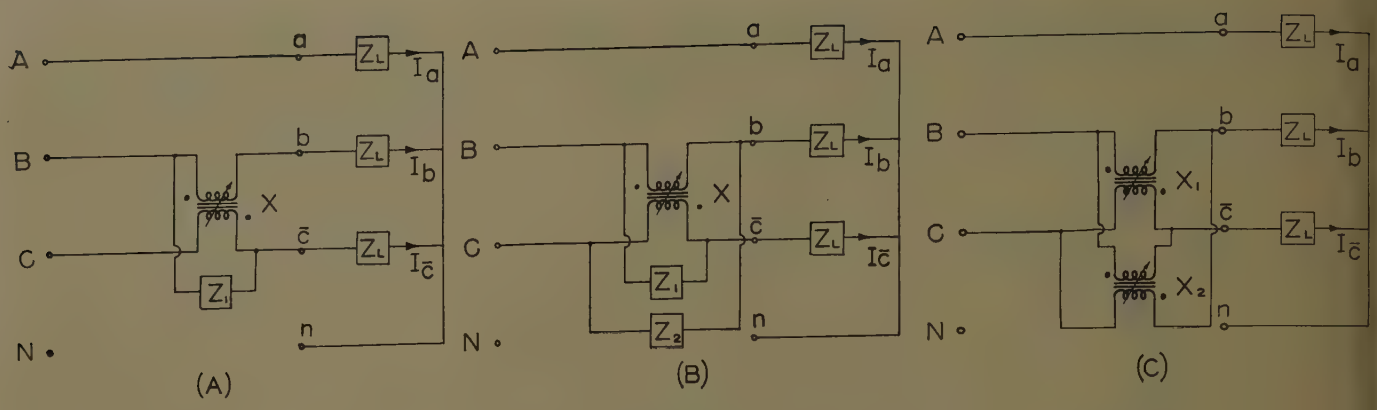


Fig. 5. Schematics of saturable-reactor interphase connections

A—Connection 2 B—Connection 3 C—Symmetrical connection 4

motor be reversible. Now, a polyphase induction motor is reversed when the phase sequence of its applied voltages is reversed. Therefore the output voltages of any magnetic power modulator which is used for driving an induction motor must, as a basic requirement, be capable of controlled phase inversion.

Several interphase combinations of two or more saturable reactors may be used to produce controllable, unbalanced voltages with the additional property of controlled phase inversion. For convenience these are classified here into two groups according to whether the interconnection involves two of the phase conductors or all three.

Reactor Interconnections Between Two Phases

Figs. 5(A), (B), and (C) show the schematic circuit arrangements of connections 2, 3, and 4 respectively. If ideal series-connected saturable reactors are used as the interconnection impedances Z_1 and Z_2 in Fig. 5(A) and (B), it is found that each of the connections 2, 3, and 4 has identical theoretical load voltage loci for the same fixed load. The locations of the load voltage vectors on these fixed loci are determined only by the various saturable reactor control currents.

The three connections 2, 3, and 4 have the additional common feature that the load phase voltage V_{an} is at all times coincident with applied phase voltage V_{AN}. No displacement of the load neutral point occurs, irrespective of the type of loading or of the nature of interconnection impedances Z_1, and Z_2.

An outline analysis of connection 2, Fig. 5(A), is given in Appendix II for the general load impedance $Z_L=R_L+jX_L$ and the general interconnection impedance $Z_1=R_1+jX_1$, used with a unity turns ratio reactor. Analyzed in terms of all three variables Z_L, Z_1 and saturable reactor reactance X, the load voltage loci are fourth power functions of forbidding complexity. When, however, a constraint is placed in turn on each of the controlling variables Z_1 and X, the load voltage loci are found to be circles similar to the circular loci of basic connection 1.

The performance of connection 2, Fig. 5(A), with reactance X infinite, is summarized by equations 22, 24, and 25, reproduced here for convenience.

$$\tan\beta=\left[X_1-R_1\frac{X_L+2X_1}{R_L+2R_1}\right]\Big/\left[R_1+X_1\frac{X_L+2X_1}{R_L+2R_1}\right] \quad (22)$$

With $R_1=0$

$$V_{Cb}=\frac{V_{CB}}{2}\left[\cos\beta-\frac{X_L}{R_L}\sin\beta\right]\angle\beta \quad (24)$$

With $X_1=0$

$$V_{Cb}=\frac{V_{CB}}{2}\left[\cos\beta+\frac{R_L}{X_L}\sin\beta\right]\angle\beta \quad (25)$$

When the interconnection impedance Z_1 in Fig. 5(A) is an ideal saturable reactor, $R_1=0$ and the voltage loci are given by equation 24. With a resistive load, $X_L=0$ and the negative part of equation 24 disappear. Voltage V_{Cb} then has a semicircular locus on the diameter $V_{CB}/2$, the location of the vector being determined by saturable reactor reactance X_1.

With X_1 infinite, connection 2 becomes identical with the basic interphase connection 1. Points $\bar{c}$ and b in triangle ABC then traverse circular loci from their positive sequence locations C and B respectively when $X=0$ to coincidence at $V_{CB}/2$ and when $X=\infty$, as shown in Fig. 3. Let reactor X now be maintained at infinity while interconnection reactor X_1 goes into saturation. Load voltages

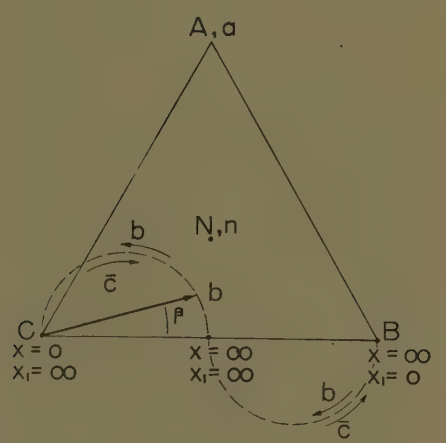


Fig. 6. Load voltage loci for connection 2, with purely resistive load, when Z_1 is a series saturable reactor

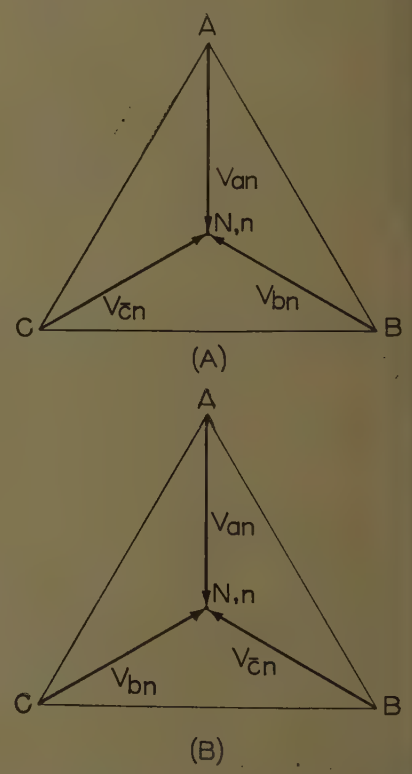


Fig. 7. Limiting load voltage conditions for connection 2

A—X=0, $X_1=\infty$
B—X=∞, $X_1=0$

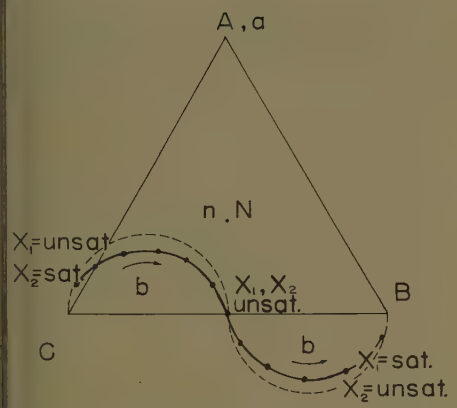


Fig. 8. Measured load voltage loci for symmetrical connection 4 with resistive load

V_{bn} and $V_{\bar{c}n}$ then proceed from coincidence at $V_{CB}/2$ around semicircle loci to coincide with applied voltages V_{CN}, V_{BN} respectively, thus forming a balanced negative-sequence set. The loci of these load voltages are shown in Fig. 6 with the two extreme conditions also shown in Fig. 7.

When this control sequence is applied to nonresistive loads, the voltage loci still perform excursions between the two limits of Fig. 7 but these loci, now described by the complete equation 24, are then of the type shown in Fig. 4. Connections 3 and 4 also produce voltage loci identical with those of Figs. 6 and 7, when appropriate control sequences are applied.

Measured Performance of Symmetrical Connection 4

Shown in Fig. 8 are the measured load voltage loci for a resistive load used with the symmetrical 2-reactor connection 4, from Fig. 5(C).

Comparison of the measured and theoretical characteristics illustrates the degree of approximation involved in the idealizing of the saturable reactors. The unsaturated reactance of the two unity turns ratio reactors were large enough to give good agreement at the phase inversion point $V_{CB}/2$. It is seen that the measured characteristics are fairly symmetrical about chords $V_{BC}/2$ and $V_{CB}/2$. Equations 24 and 25 therefore suggest that the saturable reactors are, as might be expected, behaving like R-L circuits of approximately constant power factor rather than as lossless inductors.

A minimum voltage unbalance is evident at each extreme control position in Fig. 8 owing to the nonzero saturated impedance of the reactors. This limit on the degree of voltage balance attainable may be minimized by using core materials of minimum saturated permeability. It may be almost eliminated by using fixed capacitors in series with each power winding, designed to resonate at the maximum reactor control curvant setting. However, the possibility of a complete balance of applied voltages, though desirable, is not essential for the induction motor in a variable speed drive. In many cases, because of the additional cost and bulk of capacitive compensation to eliminate a small percentage unbalance, it may be considered an unnecessary refinement.

The maximum voltage across the reactor power windings is equal to the applied line-to-line voltage for any R-L load. In Fig. 6, for instance, voltage $V_{C\bar{c}}$, V_{Bb}, $V_{B\bar{c}}$ and V_{Cb} each has its maximum values when coincident with vectors V_{CB} or V_{BC}. A capacitive load would introduce the possibility of excessive steady-state reactor voltages as well as the undesirable possibility of subharmonic ferroresonances.

A measured harmonic analysis of load voltage waveforms showed that the degree of distortion was consistent with the characteristics of Fig. 8. With both reactors unsaturated there was only 2% fifth harmonic, while at the corners B and C of triangle ABC, the most balanced condition, the waveforms contained 15% third harmonic and 5% fifth harmonic. The waveform oscillograms in this latter condition showed barely perceptible departures from the sinusoidal. At values of reactor control currents corresponding to the peaks of the semicircles in Fig. 8, a maximum harmonic content of 40% third harmonic was measured, together with much smaller quantities of other odd harmonics.

Symmetrical Components of Connections 2, 3, and 4

The torque–speed characteristic of a polyphase induction motor supplied by unbalanced applied voltages may conveniently be considered as the algebraic sum of two separate characteristics because of the positive-sequence and negative-sequence components of applied voltage acting independently. For accurate and sensitive control of both hoisting and lowering loads it is desirable that the unbalanced voltages be continuously adjustable to give any desired combination of positive-sequence and negative-sequence torque. The thermal design of the motor and the duty cycle of the drive will determine the limitation of permissible current unbalance.

The relative magnitudes of the two sequence voltages given by symmetrical connection 4 with a resistive load are shown in Fig. 9(A). The theoretical characteristic is found to consist of two elliptical sections intersecting at the phase inversion point. Also plotted in Fig. 9(A) is the symmetrical component characteristic corresponding to the measured locus in Fig. 8.

When a purely inductive load is used with symmetrical connection 4, the theoretical symmetrical component locus is the linear characteristic shown in Fig. 9(B). The theoretical locus for any R-L load will therefore lie within the shaded area in this figure.

Since connections 2, 3, and 4 have the same theoretical load voltage loci, their theoretical symmetrical component loci are also identical with the characteristics of Fig. 9. Connection 3, Fig. 5(B), has the disadvantage of requiring three saturable reactors but gives highly sensi-

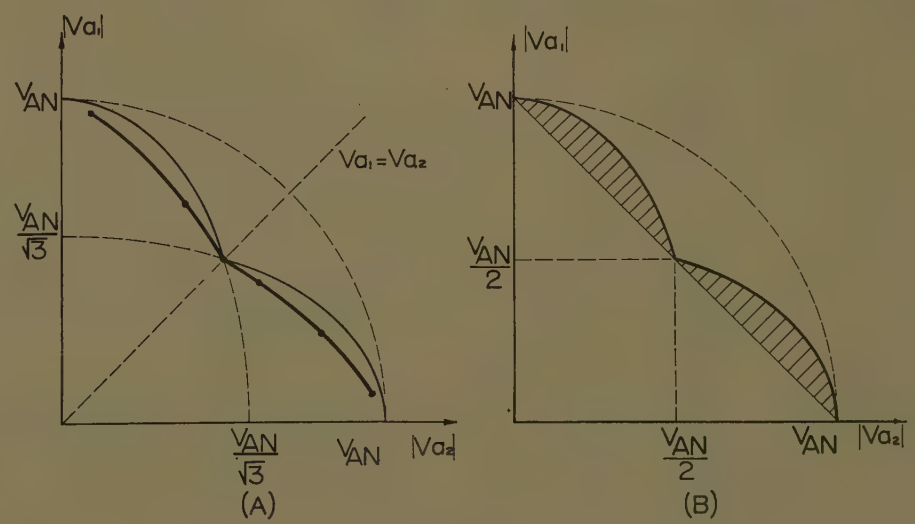


Fig. 9. Symmetrical component load voltages for connection 4

A—Measured and theoretical loci with resistive load
B—Theoretical boundary loci for resistance-inductance loads

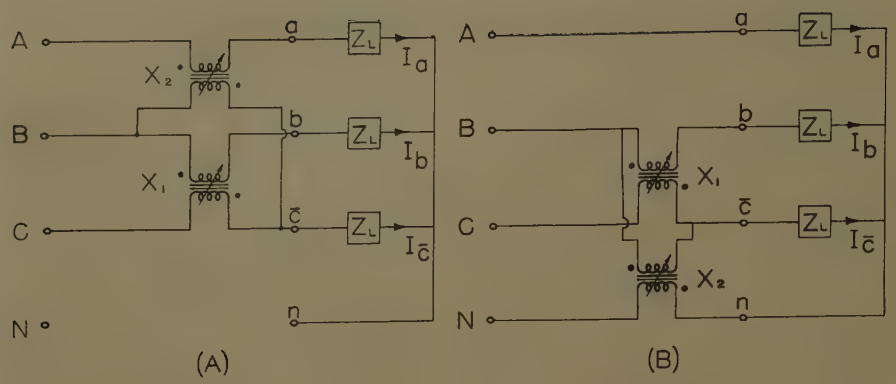


Fig. 10. Saturable reactor interphase connections

A—Cross connection 5 B—Neutral connection 6

tive control of the load voltages. Connections 2 and 4, shown in Figs. 5(A) and (C) respectively each require two reactors and appear to be equal in cost and bulk as well as identical in performance. L. R. Foote has described a scheme which is essentially that of connection 2 with a third reactor in series with line A. The symmetrical connection 4 is found, after considerable topological rearrangement, to be identical with a circuit suggested by Leonhard[5] and uses two identical reactors operating in identical modes. It has the advantage of resulting in a compact and elegant sinusoidal analysis.

Reactor Interconnections Between Three Phases

Variable reversible 3-phase voltages may also be obtained by connecting two saturable reactors, as shown in Figs. 10(A) and (B). These are classified as connections 5 and 6 and are referred to as the cross connection and the neutral connection respectively.

An analysis of cross connection 5 for the general load $Z_L = R_L + jX_L$ with two similar unity turns ratio reactors results once more in complex fourth-power equations for the load voltage loci. To reduce the analysis to a manageable complexity, reactors X_1 and X_2, in Fig. 10(A) are considered ideal, and each in turn is constrained in value. The analysis, not included in the paper, gives the following results when $X_1 = \infty$, X_2 = finite; then

$$\tan \theta_2 = \frac{3R_L}{3X_L + 8X_2} \tag{26}$$

$$V_{C\bar{b}} = \frac{3V_{NB}}{4}\left(\cos \theta_2 - \frac{X_L}{R_L} \sin \theta_2\right) \angle \theta_2 \tag{27}$$

When $X_2 = \infty$ X_1 = finite, then

$$\tan \theta_1 = \frac{3R_L}{3X_L + 8X_1} \tag{28}$$

$$V_{B\bar{b}} = V_{BQ}\left(\cos \theta_1 - \frac{X_L}{R_L} \sin \theta_1\right) \angle \theta_1 \tag{29}$$

It is evident that equations 11, 24, 27, and 29 are of identical form, differing only in the magnitude of the constant multiplier. In each case this multiplier defines the chord on which the circular load voltage loci are based. For resistive loads the negative part of the equations disappears and the chord becomes a diameter.

Fig. 11(A) shows the theoretical voltage loci of cross connection 5 for a resistive load over the control cycle of reactor X_2 while reactor X_1 is maintained at infinite impedance. The voltage vectors are shown at an arbitrary intermediate value of control current in reactor X_2. With $X_1 = \infty$ $X_2 = 0$, points b and $\bar{c}$ coincide with points C and B respectively to give a balanced negative sequence. As X_2 then increases, the load voltages describe circular loci, moving to points q when $X_2 = \infty$. The load neutral point also describes a circular locus in which the line-to-line load voltage $V_{a\bar{c}}$ is at all times constant and equal to V_{AC}.

If reactance X_2 is now maintained infinite while reactor X_1 goes into saturation, the voltages describe further circular paths from points q to points r in Fig. 11(B) which, for simplicity, shows only the loci of points a and n.

A set of measured load voltage loci for cross connection 5 are shown superposed on the complete theoretical voltage diagram in Fig. 11(C). Good agreement is evident only for the locus of phase voltage $V_{\bar{c}n}$.

Voltages V_{an} and V_{bn} exhibit a considerable discrepancy from their respective theoretical loci over the whole control cycle. It is evident that the saturable reactors are behaving nonideally and that the use of this connection as a power modulator in a close loop variable-speed drive would require consideration in terms of measured rather than ideal theoretical performance. The use of cross connection 5 with nonresistive loads is governed by equations 27 and 29.

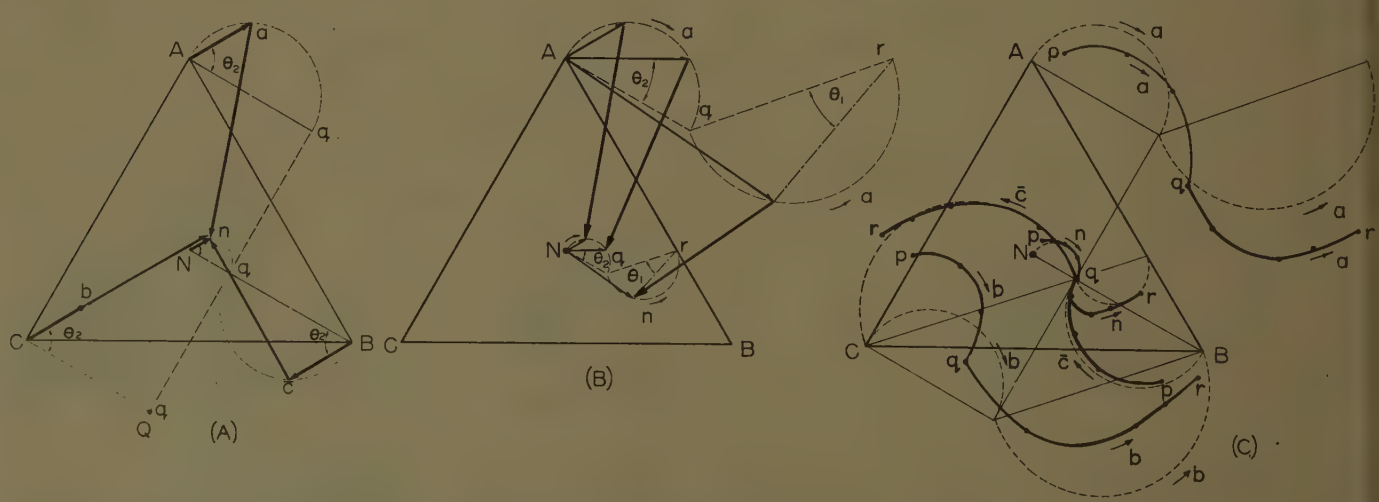


Fig. 11. Load voltage loci for cross connection 5 with resistive load

A—As X_1 varies from zero to ∞; $X_1 = \infty$ B—Same as A for $X_2 = \infty$ C—Measured and theoretical

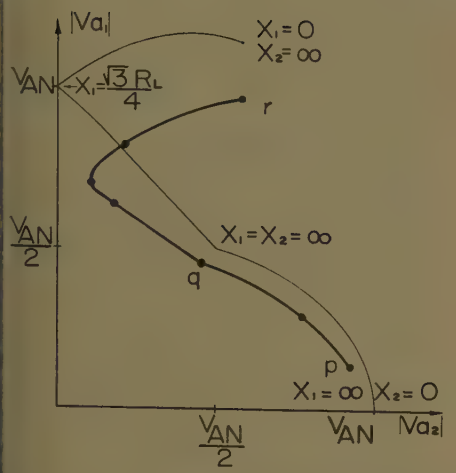


Fig. 12. Load voltage symmetrical component loci for cross connection 5 with resistive load

As with previous connections, the presence of load reactance merely changes the radii and center displacements of the circular loci, which remain based on the chords shown in Fig. 11(C).

It is found by experiment that the performance of neutral connection 6, shown in Fig. 10(B), is very similar to that of cross connection 5.

Undesirably large neutral currents are necessary to achieve certain load voltage unbalances. This connection was not found to have any advantage over the performance of cross connection 5.

Symmetrical Component Loci of Cross Connection 5

The measured and theoretical loci of the symmetrical component load voltages V_{a1} and V_{a2}, corresponding to Fig. 11(C) are shown in Fig. 12. The lower theoretical locus, where $X_1 = \infty$, is found to be elliptical, and identical to the corresponding portion of Fig. 9(A) for symmetrical connection 4. The upper theoretical portion of Fig. 12, where $X_2 = \infty$ is seen to have a break point at rated applied voltage for the case of a resistive load R_L.

The considerable divergence between the measured and theoretical results in the voltage loci of Fig. 11(C) is reproduced in the symmetrical component loci of Fig 12. With a purely inductive load the theoretical voltage loci of Fig. 11(C) would become the chords of the present semicircles. Fig. 13 shows the corresponding theoretical symmetrical component locus for cross connection 5 with an inductive load. The shaded section in Fig. 13 is thus the theoretical boundary area for R-L loads.

A measured harmonic analysis of the load voltage waveforms for cross connection 5 with a resistive load gave results consistent with the loci of Fig. 11(C). At points p for instance, where $X_1 = \infty$, $X_2 = 0$, load voltages V_{bn}, $V_{\bar{c}n}$ each contained 17% third harmonic and 6% fifth harmonic while voltage V_{an} contained only 1.0% of each. At points r, on the other hand, voltages V_{an} and $V_{\bar{c}n}$ contained 17% third and 5% fifth harmonic while voltage V_{bn} was very nearly a pure sinusoid. With both reactors unsaturated, at points q, voltages V_{an} and V_{bn} contained only 3% third harmonic.

Comparison of the symmetrical component loci of Figs. 9 and 12, for connections 4 and 5 respectively, indicate their relative usefulness as magnetic power modulators for driving polyphase induction motors. The theoretical locus of cross connection 5, Fig. 12, shows promise, since positive-sequence voltages greater than the applied voltage are obtainable.

Considering the two theoretical loci only, connection 5 appears to be more useful than the symmetrical connection 4. It is evident, however, that the measured locus of symmetrical connection 4, Fig. 8, approaches much more nearly its ideal theoretical performance than does the corresponding performance of cross connection 5, Fig. 11(B). Compared on a basis of measured performance, the symmetrical connection 4 appears to be the more useful. A final choice between the two connections would have to be made according to the quality of saturable reactors available and to the particular load voltages desired.

Conclusions

The paper has described five methods by which two saturable reactors may be interconnected between the lines of a balanced 3-phase supply to give controllable, reversible 3-phase output voltages. Several of the connections appear to be directly applicable as power modulators for driving polyphase induction motors in stepless a-c variable-speed drives.

The performance of the modulators can be predicted approximately by regarding the saturable reactors as lossless variable inductors and with the use of sinusoidal steady-state analysis.

Appendix I. Outline Analysis of Basic Connection 1

$$I_a + I_b + I_{\bar{c}} = 0 \tag{1}$$

$$V_{AB} = I_a Z_L - I_b(Z_L + jX) + I_{\bar{c}} jX \tag{2}$$

$$V_{BC} = I_b(Z_L + j2X) - I_{\bar{c}}(Z_L + j2X) \tag{3}$$

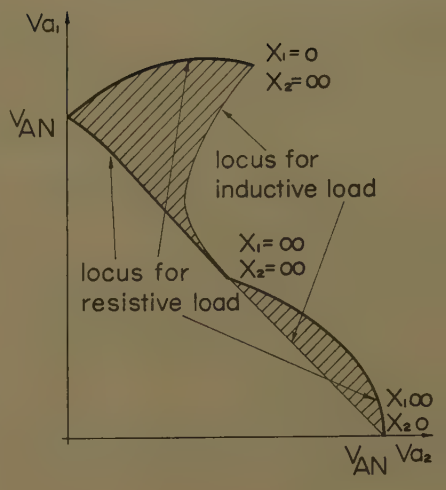


Fig. 13. Theoretical region of symmetrical components loci for R-L load with cross connection 5

Combining equation 1 and equation 2 yields

$$V_{AB} = I_a(2Z_L + jX) + I_{\bar{c}}(Z_L + j2X) \tag{4}$$

Combining equations 1 and 3:

$$V_{BC} = -I_a(Z_L + j2X) - I_{\bar{c}}2(Z_L + j2X) \tag{5}$$

Combining 4 and 5:

$$V_{an} = I_a Z_L = \frac{2V_{AB} + V_{BC}}{3} = \text{constant} \tag{6}$$

Combining 4, 5, and 6:

$$V_{\bar{c}n} = I_{\bar{c}} Z_L = \frac{V_{CB}(2Z_L + jX) + V_{BA}(Z_L + j2X)}{3(Z_L + j2x)} \tag{7}$$

Put

$$Z_L = R_L + jX_L$$
$$V_{BA} = V_{CB}(-1/2 + j\sqrt{3}/2)$$
$$V_{C\bar{c}} = V_{Cn} - V_{\bar{c}n} \tag{8}$$

Combining 7 and 8:

$$V_{C\bar{c}} = V_{CB}\left[\frac{X}{\sqrt{R_L^2 + (2X + X_L)^2}}\right] \times \underline{\left|\tan^{-1}\frac{R_L}{2X + X_L}\right.} \tag{9}$$

Let

$$\tan\alpha = \frac{R_L}{2X + X_L} \tag{10}$$

Combining 9 and 10:

$$V_{C\bar{c}} = \frac{V_{CB}}{2}\left[\cos\alpha - \frac{X_L}{R_L}\sin\alpha\right] \underline{\left|\alpha\right.} \tag{11}$$

Cartesian co-ordinates x and y of $V_{C\bar{c}}$ are:

$$\left[x - \frac{V_{CB}}{4}\right]^2 + \left[y + \frac{V_{CB}X_L}{4R_L}\right]^2 = \left[\frac{V_{CB}}{4}\right]^2 + \left[\frac{V_{CB}X_L}{4R_L}\right]^2 \tag{12}$$

Now,

$$|Z_L| = \sqrt{R_L^2 + X_L^2} \quad (13)$$

Combining 12 and 13:

$$\left[x - \frac{V_{CB}}{4}\right]^2 + \left[y + \frac{V_{CB} X_L}{4R_L}\right]^2 = \left[\frac{V_{CB}|Z_L|}{4R_L}\right]^2 \quad (14)$$

Combining 13 and 14:

$$\left[x - \frac{2X^2 V_{CB}}{4X^2 - Z_L^2}\right]^2 + y^2 = \left[\frac{V_{CB} X |Z_L|}{4X^2 - |Z_L|^2}\right]^2 \quad (15)$$

Appendix II. Outline Analysis of Connection 2, Fig. 5

$$I_a + I_b + I_c = 0 \quad (16)$$

$$V_{AB} = I_a Z_L - I_b(Z_L + jX) + I_c jX \quad (17)$$

$$V_{BC} = I_b(Z_L + j2X) - I_c Z_L - I_c j2X \quad (18)$$

$$V_{BC} = I_b jX + I_c Z_1 - I_C(Z_1 + jX) \quad (19)$$

Combining equations 16, 17, and 18 yields

$$V_{an} = I_a Z_L = \frac{2V_{AB} + V_{BC}}{3} = \text{independently of } Z_1 \quad (20)$$

If the value of X is now kept at infinity

$$V_{Cb} = V_{CB} \frac{Z_1}{Z_L + 2Z_1} \quad (21)$$

Let

$$\tan \beta = \frac{X_1 - R_1 \dfrac{X_L + 2X_1}{R_L + 2R_1}}{R_1 + X_1 \dfrac{X_L + 2X_1}{R_L + 2R_1}} \quad (22)$$

Then

$$V_{Cb} = |V_{CB}| \frac{\sqrt{R_1^2 + X_1^2}}{\sqrt{(R_L + 2R_1)^2 + (X_L + 2X_1)^2}} \underline{|\beta} \quad (23)$$

With

$$R_1 = 0$$

$$V_{Cb} = \frac{V_{CB}}{2}\left[\cos \beta - \frac{X_L}{R_L} \sin \beta\right] \underline{|\beta} \quad (24)$$

With

$$X_1 = 0$$

$$V_{Cb} = \frac{V_{CB}}{2}\left[\cos \beta + \frac{R_L}{X_L} \sin \beta\right] \underline{|\beta} \quad (25)$$

References

1. Speed Control of Induction Motors Using Saturable Reactors, **P. L. Alger, Y. H. Ku.** *AIEE Transactions*, pt. III (*Power Apparatus and Systems*), vol. 75, 1956 (Feb. 1957 section), pp. 1335–41.

2. Torque and Speed Control of Induction Motors Using Saturable Reactors, **J. F. Szablya.** *Ibid.*, vol. 77, 1958 (Feb. 1959 section), pp. 1676–82.

3. A Variable-Speed Reversible Drive Using an Induction Motor, **G. Hausen, P. P. Biringer, G. R. Slemon.** *Ibid.*, pt. III-B, vol. 78, 1959, (Feb. 1960 section), pp. 1549–54.

4. Adjustable Speed Control of A-C Motors, **L. R. Foote.** *Electrical Engineering*, vol. 78, no. 8, Aug. 1959, pp. 840–43.

5. Elements of Reactor-Controlled Reversible Induction-Motor Drives **Werner Leonhard.** *AIEE Transacsions*, pt. II (*Applications and Industry*), vol. 78, May 1959, pp. 106–15.

6. Symmetrical Components (book), **C. F. Wagner, R. D. Evans.** McGraw-Hill Book Company, Inc., New York, N. Y., 1933, chap. XVII.

Plant Identification in the Presence of Disturbances and Application to Digital Adaptive Systems

P. JOSEPH STUDENT MEMBER AIEE **J. LEWIS** ASSOCIATE MEMBER AIEE **J. TOU** MEMBER AIEE

THE NEED for an adaptive control system occurs when the parameters involved in the design of the system change in time. For small variations, conventional feedback configurations can sometimes reduce the sensitivity of the output to these changes to acceptable values, but large variations of the parameters require a more flexible system. The newer adaptive point of view in designing a control system can, in general, allow for changes in input signal characteristics, plant characteristics, and disturbance characteristics with a minimum amount of *a priori* knowledge. It appears that a general distinction has been made between adaptive controls[1,2] and optimalizing (optimizing) controls[3,4] although the characteristics are not always distinct. A recent note[5] has suggested a possible classification according to whether there is an original feedback loop or not.

Most methods of designing control system compensators assume that the open-loop plant transfer function is known. If the open-loop plant transfer function could be determined, one of many compensator design methods could be mechanized to yield a plant-adaptive control system. A major problem in this approach to adaptive control, therefore, is the determination of the transfer function: the identification problem.

Any realistic treatment of the identification problem must include the effects of disturbances upon the accuracy of the identification. The presence of disturbances prevents an exact identification in a finite time and, therefore, a compromise between accuracy and speed of identification is necessary.

In order adequately to describe a plant by a transfer function, the time variation of the plant parameters must be slow in relation to the significant system time constants. The method of identification used here actually assumes that the parameters are constants between periods of adaptation. The relationship between rate of change of plant parameters and period of adaptation depends on many factors such as the identification accuracy required and the ratio of signal to output disturbance. For example, assume that a preselected identification accuracy of 1% is chosen and that the signal/disturbance ratio is such that the time required for this identification is 10 sampling periods. If a plant parameter is drifting at a rate of less than 0.1% per sampling period, it can be assumed to be constant to within the identification accuracy during the identification interval.

In carrying out the identification, it is assumed that both the number of poles and the number of zeros of the transfer function are known. If only an upper bound on each of these numbers is known, it is sufficient, although it is desirable to keep these bounds small since the computations required increase rapidly.

Identification

Consider the system of Fig. 1. $G(z)$ is the open-loop plant transfer function and

Paper 61-73, recommended by the AIEE Feedback Control Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 29–February 3, 1961. Manuscript submitted October 25, 1960; made available for printing December 15, 1960.

P. Joseph, J. Lewis, and J. Tou are with the School of Electrical Engineering, Purdue University, Lafayette, Ind.

This work has been supported by the Office of Naval Research under Contract *Nonr-1100(18)*.

$$(z)=\frac{V(z)}{M(z)} \tag{1}$$

$V(z)$ is a disturbance and the output, $(z)=V(z)+U(z)$. The input sequence $R(z)$ and it is assumed that the input nd the disturbance are independent. t is assumed that the past values of m re known exactly since they are generated y the discrete filter which will be realized y a digital computer. It is also assumed hat the past values of r and c can be easured. Measurement errors in r will ot adversely affect the adaptation rocess and measurement errors in c will ave an effect similar to that of u. herefore, they will not be considered. is assumed to be unmeasurable.

If

$U(z)=0$

hen

$$G(z)=\frac{C(z)}{M(z)} \tag{2}$$

n reference 6 $G(z)$ is assumed to have he form

$$G(z)=\frac{\sum_{j=0}^{\xi} a_j z^{-j}}{1+\sum_{j=1}^{\eta} b_j z^{-j}} \tag{3}$$

where ξ and η are known constants or at least an upper bound of ξ and η are known. By using equations 2 and 3, a set of simultaneous linear equations in the $\xi+\eta+1$ unknowns, a_j and b_j, can be derived. Any measurement error in c can be reduced by the use of least squares smoothing.[6] The success of the smoothing depends upon m being independent of the measurement error.[7]

When $U(z)\neq 0$, one might attempt to estimate $G(z)$ by $\hat{G}(z)$ where

$$\hat{G}(z)=\frac{C(z)}{M(z)} \tag{4}$$

$$\hat{G}(z)=\frac{V(z)}{M(z)}+\frac{U(z)}{M(z)}=G(z)+\frac{U(z)}{M(z)} \tag{5}$$

The use of least squares smoothing does not help in this case, since m is not independent of u. This is a result of the closed-loop nature of the system.

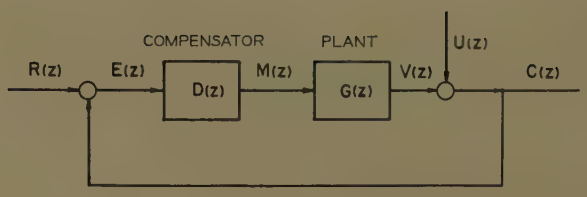
As a starting point, an alternative to equation 1 is

$$G(z)=\frac{\Phi_{xv}(z)}{\Phi_{xm}(z)} \tag{6}$$

where $\Phi_{xm}(z)$ is the correlation sequence between x and m defined by the relations

$$\phi_{xm}(k)=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{i=-N}^{N} x_i m_{i+k} \tag{7}$$

Fig. 1. System diagram



and

$$\Phi_{xm}(z)=\sum_{k=-\infty}^{\infty}\phi_{xm}(k)z^{-k} \tag{8}$$

and where x is any signal which is correlated with m.

When $U(z)\neq 0$, one might attempt to estimate $G(z)$ by $\hat{G}(z)$ where

$$\hat{G}(z)=\frac{\Phi_{xc}(z)}{\Phi_{xm}(z)} \tag{9}$$

$$\hat{G}(z)=\frac{\Phi_{xv}(z)+\Phi_{xu}(z)}{\Phi_{xm}(z)}=G(z)+\frac{\Phi_{xu}(z)}{\Phi_{xm}(z)} \tag{10}$$

If x is chosen to be a signal that is not correlated with u, $\hat{G}(z)=G(z)$. The obvious choice for x is r. Then

$$G(z)=\frac{\Phi_{rc}(z)}{\Phi_{rm}(z)}\equiv\hat{G}(z) \tag{11}$$

The solution of the identification problem cannot, however, be left here. The evaluation of $\Phi_{rc}(z)$ and $\Phi_{rm}(z)$ requires taking an infinite limit. This is not only impractical actually to implement, but would utilize information from the distant past that would be out of date if the plant were time varying. If only recent information is to be utilized, an exact determination of $G(z)$ is impossible. However, a good approximation is available. The following definitions are, therefore, made:

$$\theta_{rm}(k,N,I)=\sum_{i=I-N}^{I} r_i m_{i+k} \tag{12}$$

$$\textcircled{H}_{rm}(z,N,I)=\sum_{k=-\infty}^{\infty}\theta_{rm}(k,N,I)z^{-k} \tag{13}$$

Reference 8 contains a proof of equation 6, and a completely analogous argument shows that

$$G(z)=\frac{\textcircled{H}_{rv}(z,N,I)}{\textcircled{H}_{rm}(z,N,I)} \tag{14}$$

$G(z)$ will be estimated by $\hat{G}(z)$ where

$$\hat{G}(z)=\frac{\textcircled{H}_{rc}(z,N,I)}{\textcircled{H}_{rm}(z,N,I)} \tag{15}$$

If

$U(z)=0$

then

$G(z)=\hat{G}(z)$

If

$U(z)\neq 0$

then

$$\hat{G}(z)=\frac{\textcircled{H}_{rv}(z,N,I)+\textcircled{H}_{ru}(z,N,I)}{\textcircled{H}_{rm}(z,N,I)}=G(z)+\frac{\textcircled{H}_{ru}(z,N,I)}{\textcircled{H}_{rm}(z,N,I)} \tag{16}$$

Since

$$\lim_{N\to\infty}\frac{1}{N+1}\textcircled{H}_{ru}(z,N,I)=\Phi_{ru}(z)=0$$

in the limit as N goes to infinity

$G(z)=\hat{G}(z)$

Before discussing the closeness of the estimation when N is finite, equation 15 will be manipulated into a form which can be used to solve for the $\hat{a}_j$'s and $\hat{b}_j$'s which determine $\hat{G}(z)$. $\hat{G}(z)$, the estimate of $G(z)$, will be expressed in a form similar to equation 3

$$\hat{G}(z)=\frac{\sum_{j=0}^{\xi}\hat{a}_j z^{-j}}{1+\sum_{j=1}^{\xi}\hat{b}_j z^{-j}} \tag{17}$$

$$\hat{G}(z)=\frac{\sum_{j=0}^{\xi}\hat{a}_j z^{-j}}{1+\sum_{j=1}^{\xi}\hat{b}j z^{-j}}=\frac{\sum_{k=-\infty}^{\infty}\theta_{rc}(k,N,I)z^{-k}}{\sum_{k=-\infty}^{\infty}\theta_{rm}(k,N,I)z^{-k}} \tag{18}$$

Cross multiplying, letting $k=-j-p$, and equating coefficients of like powers of z, yields

$$\sum_{j=0}^{\xi}\hat{a}_j\theta_{rm}(-j-p,N,I)-\sum_{j=1}^{\eta}\hat{b}_j\theta_{rc}(-j-p,N,I)=\theta_{rc}(-p,N,I) \tag{19}$$

By selecting $\xi+\eta+1$ distinct values of p, $\xi+\eta+1$ linear simultaneous equations can be obtained from equation 19. These equations can then be solved for the $\hat{a}_j$'s and the $\hat{b}_j$'s which determine $\hat{G}(z)$.

Returning to the problem of the accuracy of the estimation of $G(z)$ by $\hat{G}(z)$, or more precisely, the estimation of the a_j's and b_j's by the $\hat{a}_j$'s and $\hat{b}_j$'s, it is found by examination of equations 14, 15, 16, and 19 that the a_j's and b_j's satisfy

$$\sum_{j=0}^{\xi}a_j\theta_{rm}(-j-p,N,I)-\sum_{j=1}^{\eta}b_j\theta_{rv}(-j-p,N,I)=\theta_{rv}(-p,N,I) \tag{20}$$

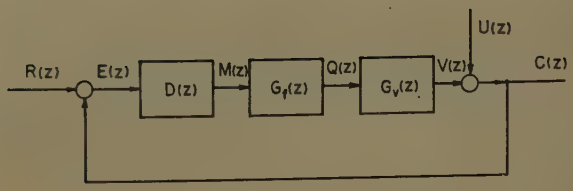


Fig. 2. System diagram

$$\sum_{j=0}^{\xi} a_j\theta_{rm}(-j-p, N, I)-\sum_{j=1}^{\eta} b_j[\theta_{rc}(-j-p, N, I)-\theta_{ru}(-j-p, N, I)]=\theta_{rc}(-p, N, I)-\theta_{ru}(-p, N, I) \quad (21)$$

Comparing equations 21 and 19 shows that the coefficients of b_j and $\hat{b}_j$ differ by $\theta_{ru}(-j-p,N,I)$. Clearly

$$E[\theta_{ru}(k, N, I)]=0 \quad (22)$$

and in Appendix I it is shown that

$$\sigma_\theta^2=E\{[\theta_{ru}(k, N, I)]^2\}=\sum_{i=I-N}^{I}\sum_{j=I-N}^{I}\phi_{rr}(i-j)\phi_{uu}(i-j) \quad (23)$$

and if u consists of white noise with variance σ_u^2.

$$\sigma_\theta^2=(N+1)\phi_{rr}(0)\sigma_u^2 \quad (24)$$

Note that the square of the mean value of $\theta_{rc}(k,N,I)$ is proportional to $(N+1)^2$ and, therefore, the percentage mean square difference between the coefficients of equations 19 and 21 decreases as $1/N+1$.

Using a white noise disturbance, as an interesting special case, the following relations are also developed in Appendix I. This development contains several approximations.

If

$$\Delta^2>K^2\theta_{rr}(O, N, I)\sigma_u^2\sum_{j,\,i}\Delta_{ij}^2 \quad (25)$$

then the 1σ error in the estimation of the a_j's (or b_j's) is

$$1\sigma_{\text{error}}=\{E[(a_j-\hat{a}_j)^2]\}^{1/2}<\frac{1}{K}(a_j^2+1)^{1/2} \quad (26)$$

where Δ is the determinant of the coefficients of the simultaneous equations and where the Δ_{ij}'s are the cofactors of the coefficients of the b_j's and K is a design parameter.

The method of calculating the $\theta_{rm}(k, N,I)$ in real time remains to be described. First, if the θ's are to be evaluated at the nth sampling instant, examination of equation 12 shows that $I\leq n$ and $I\leq n-k$ since the values of the samples which will occur in the future are unknown. In order to utilize the most recent available information, the equality will be chosen

$$I=n-k_{\max} \quad (27)$$

where $-k_{\max}$ equals the most negative value of p that is selected. Suppose that the lower limit of the summation in equation 12 is $n_0-k_{\max}$ then $I-N=n_0-k_{\max}$ or $N=I-n_0+k_{\max}$.

$$N=n-n_0 \quad (28)$$

Note also that

$$\theta_{rm}(k, n-n_0, n-k_{\max})=\theta_{rm}(k, n-1-n_0, n-1-k)_{\max}+r_{n-k_{\max}}m_{n+k-k_{\max}} \quad (29)$$

Starting with

$$\theta_{rm}(k, 0, n_0-k_{\max})=r_{n_0-k_{\max}}m_{n_0-k_{\max}+k} \quad (30)$$

the θ's can be calculated using equation 29 which requires one addition and one multiplication, per θ, per sampling instant. As time, n, goes on, N increases.

A result similar to equation 19 also can be obtained for continuous-data systems and is shown in Appendix II.

Application to an Adaptive System

The method of identifying the plant transfer function can be applied to adaptive control in a number of ways. The control or command signal can be obtained as the output of a discrete compensator which is realized on a digital computer.

The compensator must be changed in accordance with changes in the plant (or signal or disturbance) parameters and at least two general approaches to the adjustment of the compensator are possible. In one approach, the compensator is redesigned each time the system "adapts." The redesign is carried out for some preselected optimum type of compensation so that the performance criterion is inherent in the compensator design. One alternative procedure is to choose a performance criterion and calculate it from an identification of the closed-loop system. The compensator is then adjusted to give an extreme value of the performance criterion. This requires the mechanization of an optimalizing loop which is inherently slower than the first method because it utilizes a searching technique. Either the redesign or the adjustment can be done periodically or according to some plan.

The accuracy of the identification is determined from equations 25 and 26. A knowledge of the maximum tolerable 1σ errors in the a_j's and b_j's determine K by equation 26 and the plan is then to check equation 25 each sampling instant. When equation 25 is satisfied, the equations derived from equation 19 will be used to determine $G(z)$. Since the expected value of the left-hand side of 25 is proportional to $N^{2(\xi+\eta+1)}$, while that of the right-hand side is proportional to $N^{2(\xi+\eta)+1}$, equation 25 is bound to be satisfied eventually unless $\lim_{N\to\infty}\Delta=0$. To avoid this, the values of p used in equation 19 must be carefully selected to obtain a nonsingular set of equations. This is demonstrated in the example. After $G(z)$ has been calculated and the compensator redesigned, the θ's are calculated anew using equations 29 and 30 with n_1 replacing n_0 and the process repeats.

If the maximum tolerable 1σ errors are not known *a priori*, the parameter K could be set by an optimalizing loop, or in the words of a recent paper, by "the supervisor's supervisor."[9]

A large reduction in the number of simultaneous equations that must be solved is possible if $G(z)$ can be written as $G_f(z)\ G_v(z)$ where $G_f(z)$ is fixed and well-known and $G_v(z)$ is varying and unknown. The system can then be redrawn as in Fig. 2 where $Q(z)=G_f(z)M(z)$ and is easily calculated.

Writing

$$G_v(z)=\frac{\sum_{j=0}^{\xi'}A_jz^{-j}}{1+\sum_{j=1}^{\eta'}B_jz^{-j}} \quad (31)$$

and similar to equation 19

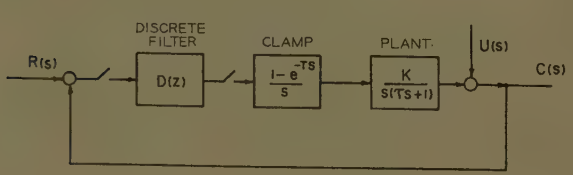


Fig. 3 (left). System of the example

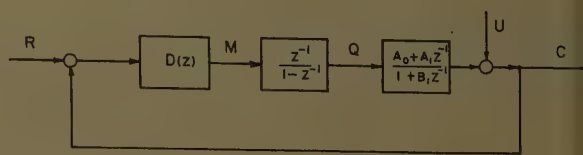


Fig 4 (right). Discrete system

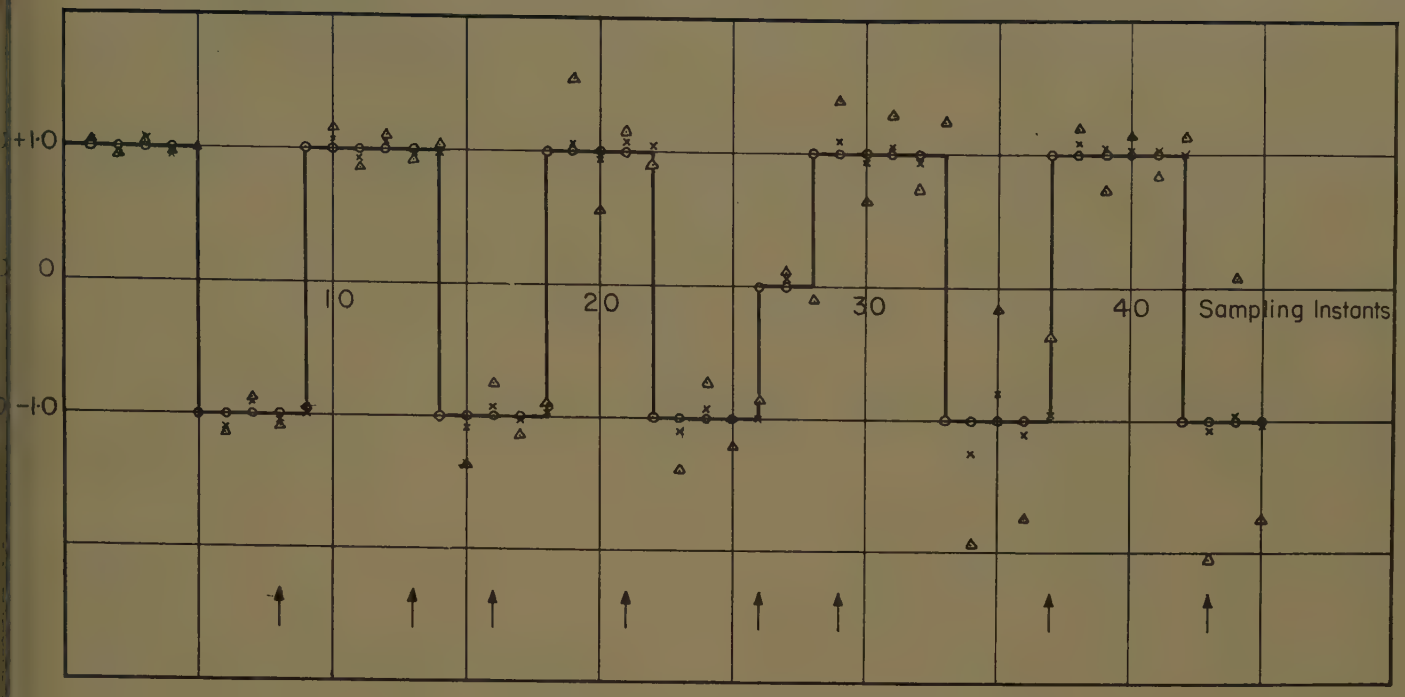


Fig. 5. Time responses for system with changing plant-gain constant

Initial $A_0=10$
Rate of change of $A_0=0.1$/sample
↑ Adaptation instants
⊙ Input
× Output with adaptation
△ Output without adaptation

$$\sum_{j=0}^{\xi'} A_j\theta_{rq}(-j-p, N, I)- \sum_{j=1}^{\eta'} B_j\theta_{rc}(-j-p, N, I)=\theta_{rc}(-p, N, I) \quad (32)$$

$G_v(z)$ and hence $G(z)$ can be determined by solving only $\xi'+\eta'+1$ equations, where $\xi'\leq\xi$ and $\eta'\leq\eta$.

Although equations 32 and 19 utilize correlation functions with r, it would also be satisfactory to correlate with a component of r. It is usually preferable to use the nonconstant components of r only.

If the disturbance does not occur at the output but elsewhere in the plant, the block diagram could still be manipulated into the form of Fig. 2 where $U(z)$ would now be an equivalent disturbance. In this instance $U(z)$ would probably not be white noise. In many situations, such as position control with torque disturbance due to a steady wind or flow control with a disturbance due to leakage, the disturbance resembles a deterministic signal more than it does white noise. If this is the case, σ_θ^2 will be much larger than predicted by equation 24. However, if a white noise test signal ρ is added to r and all correlations are performed with ρ instead of r, then

$$\sigma_\theta^2=(N+1)\phi_{\rho\rho}(0)\sigma_u^2 \quad (33)$$

This is because equation 23 is symmetrical with respect to ρ and u. Then all equations and developments hold if r is replaced by ρ. The introduction of ρ upsets the system; however, it can be made very small if long periods between adaptations are tolerable.

Example

Consider as an example the system of Fig. 3. The completely discrete version of this system is shown in Fig. 4.

$$G_v(z)=\frac{A_0+A_1z^{-1}}{1+B_1z^{-1}}$$

$$\xi'=\eta'=1 \quad (34)$$

There are three unknowns and an entire set of equations is available from equation 32. It is necessary to select three of them. Consider the equation for $p=0$.

$$\theta_{rq}(0, N, I)\hat{A}_0+\theta_{rq}(-1, N, I)_{A1}- \theta_{rc}(-1, N, I)\hat{B}_1=\theta_{rc}(0, N, I) \quad (35)$$

Suppose r consists of white noise, then due to the unit delay in the system $\varphi_{rq}(0)$, $\varphi_{rq}(-1)$, $\varphi_{rc}(0)$, and $\varphi_{rc}(-1)$ equal zero. Thus if equation 35 is selected as an equation $\lim_{N\to\infty}\Delta=0$. Thus the equations for $p=0$ and similarly all $p>0$ are rejected. Since $-k_{\max}$ is the most negative value of p selected and

$$I=n-k_{\max} \quad (27)$$

it is desirable to choose the least negative values of p possible in order to utilize the most recent information. Thus $p=-1, -2, -3$ are selected. Then

$$\theta_{rq}(1, , N,I)\hat{A}_0+\theta_{rq}(0, N, I)\hat{A}_1- \theta_{rc}(0, N, I)\hat{B}_1=\theta_{rc}(1, N, I) \quad (36)$$

$$\theta_{rq}(2, N, I)\hat{A}_0+\theta_{rq}(1, N, I)\hat{A}_1- \theta_{rc}(1, N, I)\hat{B}_1=\theta_{rc}(2, N, I) \quad (37)$$

$$\theta_{rq}(3, N, I)\hat{A}_0+\theta_{rq}(2, N, I)\hat{A}_1- \theta_{rc}(2, N, I)\hat{B}_1=\theta_{rc}(3, N, I) \quad (38)$$

These equations are solved for $\hat{A}_0$, $\hat{A}_1$, and $\hat{B}_1$ when equation 25 is satisfied.

To illustrate the design of the compensator, assume deadbeat response to a step input is desired. If all the poles and zeros of $G_v(z)$ are inside the unit circle, i.e., $|A_1/A_0|<1$, $|B_1|<1$, a straightforward application of the design procedure for deadbeat systems (reference 8, pp. 502–12) yields the following result:

$$D(z)=\frac{\alpha_0+\alpha_1z^{-1}}{1+\beta_1z^{-1}} \quad (39)$$

where

$$\alpha_0=1/A_0$$
$$\alpha_1=B_1/A_0$$
$$\beta_1=A_1/A_0 \quad (40)$$

Values for α_0, α_1, and β_1 can be found similarly for the three other possible pole and zero configurations. After A_0, A_1, B_1 are determined from equations 36, 37, and 38, the compensator, $D(z)$, is designed automatically in the computer with the aid of equation 40 and similar relations.

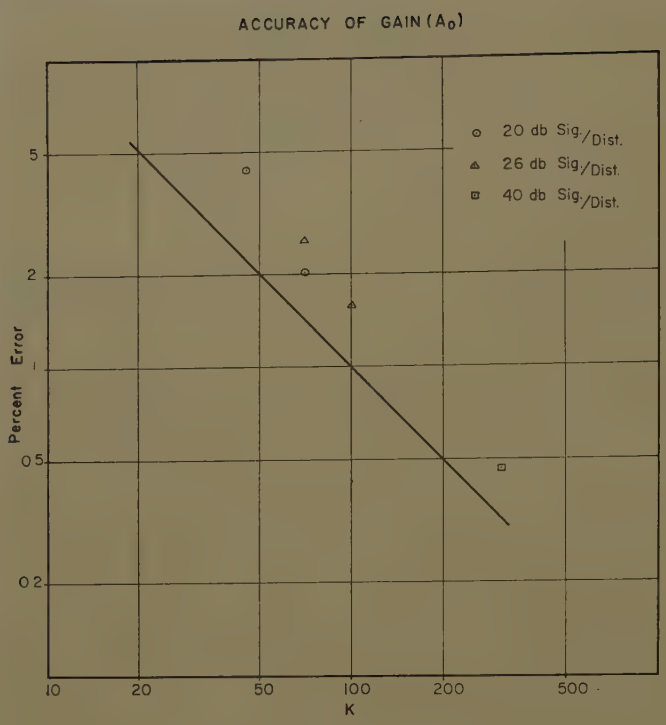


Fig. 6. Accuracy of gain (A_0)

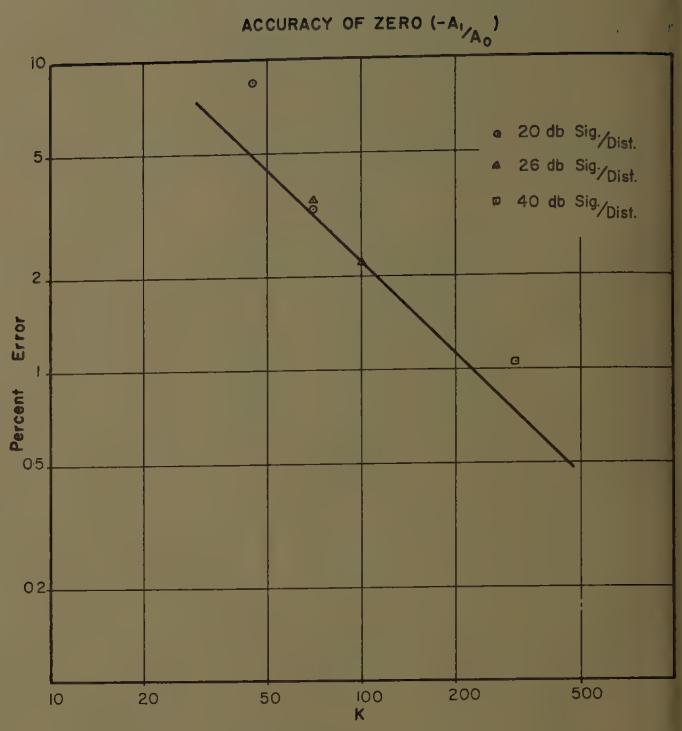


Fig. 7. Accuracy of zero ($-A_1/A_0$)

Computer Simulation

The system of the example was simulated on a digital computer. An input of the form which is shown in Fig. 5 and a white Gaussian disturbance were used. Figs. 6, 7, and 8 show the experimental 1σ errors obtained in the determination of the gain and the zero and pole locations for various fixed values of K and signal–disturbance ratio. The values of 1σ errors predicted from equation 26 are shown by the straight lines. About ten data points were averaged to determine each point on the graphs and the trend of the errors to be inversely proportional to K and to be independent of signal–disturbance ratio is demonstrated. The ability of equation 26 to estimate the errors to within a factor of 2, justifies the approximations that were made.

The period between adaptations, N, depends in a complicated manner on the shape of the input, on $D(z)$, and on $G(z)$. However, in order that the reader may form an estimate of the adaptation period, it may be noted that under the particular conditions used for the computer runs with $K=100$ and with a signal–disturbance ratio of 26 db (decibels) the adaptation time varied from 25 to 28 sampling periods. If one notes that the left-hand side of equation 25 is proportional to $(\sigma_r^2 N)^{2(\xi+\eta+1)}$ and the right-hand side to $\sigma_r^2 K^2(\sigma_r^2 N)^{2(\xi+\eta)+1}$, we see that the period between adaptations is proportional to $K^2\sigma_u^2/\sigma_r^2$. Since the errors are proportional to $1/K$, the compromise between speed and accuracy of the identification is clear.

An example of the performance of the adaptive system when the plant has a time varying parameter was obtained by allowing A_0 in the transfer function

$$G_0(z)=\frac{A_0+A_1z^{-1}}{1+B_1z^{-1}}$$

to increase at a steady rate of 0.1 per sample. The initial value of A_0 was 10 which means that the parameter was changing about 1% per sample. The rectangular wave used previously was again used as an input. The response of the adaptive system is shown in Fig. 5 and the response of the system when the compensator was left at its original correct value, i.e. a nonadaptive system, is also shown. It can be seen that the adaptive system continued to have a

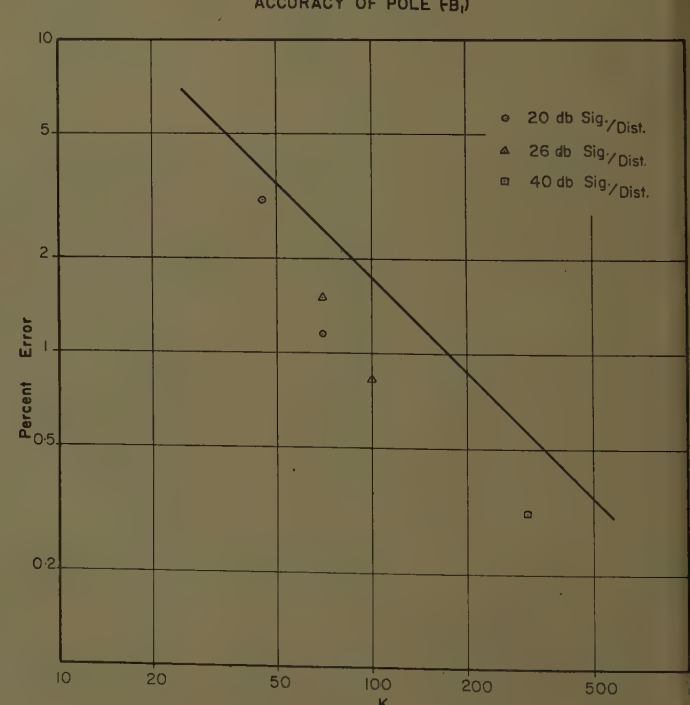


Fig. 8. Accuracy of pole ($-B_1$)

atisfactory performance whereas the onadaptive system became unsatisctory after about 20 sampling periods. he slight deviations from correct deadeat response in the adaptive system are ue to the disturbance in the output hich causes errors in the identification s well as actual disturbances to the sysem; no attempt was made to compensate o minimize disturbance effects in this ase. Some deviation from deadbeat esponse is also due to the fact that nce the parameters of the plant are contantly changing and the compensator is lways designed on the basis of past inormation, a small error in the compensaor results.

Conclusions

A method for identifying a linear plant n the presence of an output disturbance as been presented. An approximate nalysis has shown that the per-cent error n identification can be made proportional o a parameter $1/K$ while the measurement interval is proportional to K^2. imulation of a system utilizing this ethod of identification provided exerimental confirmation of these relationhips.

Appendix I. Estimate of Error in Identification

$$\mathcal{E}[\theta_{ru}(k, N, I)] = E\left[\sum_{i=-N}^{I} r_i w_{i+k}\right] = E[r]E[u](N+1) \quad (41)$$

ince r and u are assumed independentf either r or u contains no constant comonent, equation 41 equals zero. This ssumption will be made. The variance f $\theta_{ru}(k, N, I)$ is

$$\sigma_\theta^2 = E([\theta_{ru}(k, N, I)]^2) = E\left(\sum_{i=I-N}^{I}\sum_{j=I-N}^{I} r_i r_j u_{j+k} u_{i+k}\right) \quad (42)$$

Due to the independence of r and u

$$\sigma_\theta^2 = \sum_{i=I-N}^{I}\sum_{j=I-N}^{I} \phi_{rr}(i-j)\phi_{uu}(i-j) \quad (43)$$

f u consists of white noise, i.e.,

$$\phi_{uu}(0) = \sigma_u^2$$

nd

$$\phi_{uu}(k) = 0$$

or $k \neq 0$ then equation 43 becomes

$$\sigma_\theta^2 = (N+1)\phi_{rr}(0)\sigma_u^2 \quad (44)$$

ince

$$\mathcal{E}[\theta_{rr}(0, N, I)] = (N+1)\phi_{rr}(0)$$

n estimate of σ_θ^2 is

$$\sigma_\theta^2 = \theta_{rr}(0, N, I)\sigma_u^2 \quad (45)$$

Consider the set of equations

$$\begin{bmatrix} \alpha_{11} \dots \alpha_{1n} \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \alpha_{n1} \dots \alpha_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} \quad (46)$$

By Cramer's rule

$$x_1 = \frac{\begin{vmatrix} y_1 & \alpha_{12} \dots \alpha_{1n} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ y_n & \alpha_{n2} \dots \alpha_{nn} \end{vmatrix}}{\begin{vmatrix} \alpha_{11} \dots \alpha_{1n} \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \alpha_{n1} \dots \alpha_{nn} \end{vmatrix}} = \frac{D_1}{\Delta} \quad (47)$$

If the α_{ij}'s in equation 46 are replaced by $\alpha_{ij}+(\delta\alpha)_{ij}$ where $(\delta\alpha) \ll \alpha_{ij}$, the errors introduced into x_1 are of interest here. Consider the error introduced into the denominator Δ.

$$\frac{\partial \Delta}{\partial \alpha_{ij}} = \Delta_{ij} = \text{the cofactor of } \alpha_{ij} \quad (48)$$

The error introduced into Δ by $(\delta\alpha_{ij})$, is, therefore, $\Delta_{ij}(\delta\alpha)_{ij}$. If all of the $(\delta\alpha)_{ij}$'s are independent statistical quantities with standard deviation, σ_{ij}, the variance of the error introduced into Δ will be

$$\sigma_\Delta^2 = E[(\delta\Delta)^2] = \sum_i \sum_j \Delta_{ij}^2 \sigma_{ij}^2 \quad (49)$$

The assumption of the independence of the $(\delta\alpha)_{ij}$'s is not strictly applicable to the estimation of the errors in the solution of equation 19 but an approximation is all that is desired.

Since

$$\frac{D_1}{\Delta + \delta\Delta}$$

is approximately equal to

$$\frac{D_1}{\Delta} + \frac{D_1}{\Delta^2}\delta\Delta$$

the error in x_1 due to $\delta\Delta$ is

$$x_1 \frac{\delta\Delta}{\Delta}$$

Continuing the approximation assume

$$\sigma_\Delta \cong \sigma_{D1}$$

then the approximate 1σ error in x_1 is

$$[E(\delta x_1)^2]^{1/2} = \frac{1}{\Delta}\,\sigma_\Delta (x_1^2+1)^{1/2} = \sigma_{x1} \quad (50)$$

If

$$\Delta^2 < K^2 \sigma_\Delta^2$$

then

$$\sigma_{x1} < \frac{1}{K}(x_1^2+1)^{1/2} \quad (51)$$

where the condition can be rewritten

$$\Delta^2 < K^2 \sum_i \sum_j \Delta_{ij}^2 \sigma_{ij}^2 \quad (52)$$

Appendix II. Identification for Continuous Data Systems

In order to measure the continuous transfer function, $G_v(s)$, of a plant, it will be assumed that $G_v(s)$ can be written as

$$G_v(s) = \frac{\sum_{j=0}^{\xi} a_j s^j}{1 + \sum_{j=1}^{\eta} b_j s^j} \quad (53)$$

where ξ and η are known. In writing equation 53, it is assumed that $G_v(s)$ contains no integrations. Equation 53 can be rewritten as

$$G_v(s) = \frac{\sum_{j=0}^{\xi} a_j s^{j-\eta}}{s^{-\eta} + \sum_{j=1}^{\eta} b_j s^{j-\eta}} \quad (54)$$

Since

$$G_v(s) = \frac{\Phi_{rc}(s)}{\Phi_{rq}(s)} \quad (55)$$

$$\sum_{j=0}^{\xi} a_j s^{j-\eta}\Phi_{rq}(s) = s^{-\eta}\Phi_{rc}(s) + \sum_{j=1}^{\eta} b_j s^{j-\eta}\Phi_{rc}(s) \quad (56)$$

Note that the inverse Fourier transform of

$$s^{-k}\Phi_{rq}(s)$$

is

$$\phi_{rqk}(\tau)$$

where

$$q_k(t) = \int_\infty^t \dots \int_\infty^t q(t)dt^k \text{ with } k \text{ integrations}$$

Thus the inverse transform of equation 56 is

$$\sum_{j=0}^{\xi} a_j \phi_{rq\,\eta-j}(\tau) - \sum_{j=1}^{\eta} b_j \phi_{rc\,\eta-j}(\tau) = \phi_{rc\eta}(\tau) \quad (57)$$

By selecting $\xi+\eta+1$ distinct values of τ, $\xi+\eta+1$ equations can be obtained. This method requires 2η integrators to calculate the q_k's and c_k's plus $(\xi+\eta+1)^2$ multipliers and integrators to perform the correlations. In addition, facilities to solve the simultaneous equations must be provided.

References

1. A Survey of Adaptive Control Systems, J. A. Aseltine, et al. *Transactions*, Professional Group on Automatic Control, Institute of Radio Engineers, New York, N. Y., vol. AC-6, Dec. 1958, p. 102.

2. Trends in Adaptive Control Systems, J. G. Truxal. *Proceedings*, National Electronics Conference, Chicago, Ill., 1959, p. 1.

3. Engineering Cybernetics (book), H. S.

Tsien. McGraw-Hill Book Company, Inc., New York, N. Y., 1954, chap. 15.

4. Optimizing Control Systems, R. L. Cosgriff, R. A. Emerling. *AIEE Transactions*, pt. II (*Applications and Industry*), vol. 77, Mar. 1958, pp. 13–16.

5. Adaptive and Optimalizing Control Systems, P. Eykhoff. *Transactions*, Professional Group on Automatic Control, Institute of Radio Engineers, vol. AC-5, no. 2, June, 1960, p. 148.

6. Design of a Self-Optimizing Control System, R. E. Kalman. *Transactions*, American Society of Mechanical Engineers, New York, N. Y., Feb. 1958, p. 468.

7. A Study of Digital Adaptive Control Systems, J. Tou, P. Joseph, J. Lewis. *Report no. 1, Contract Nonr-1100(8)*, Office of Naval Research, Washington, D. C., Setp. 1960.

8. Digital and Sampled-Data Control Systems (book), J. Tou. McGraw-Hill Book Company, Inc., 1959.

9. Adaptive Sampled Data Systems, B. Wildrow, Butterworth's Scientific Publications, London, England.

Studies of Electrodynamic Forces Occurring at Electrical Contacts

A. C. SNOWDON
MEMBER AIEE

KNOWLEDGE of the magnitude of the electrodynamic force occurring at closed electric contacts carrying high current is important to the design of reliable contacting devices. Past work in this area by Holm[1] has shown that current passing through a contact surface is constricted to flow through very small contacting spots, and that this constricted current flow causes a repulsion force to act on the contacts. Further contributions to the theory and the calculation of this force have been made through studies at the Cutler-Hammer Laboratories in Milwaukee, Wis.,[2] the progress of which is reported in this paper.

The following analysis derives equations to calculate the magnitude of the repulsion force just mentioned. It will be shown that the force depends on contact pressure, diameter, material hardness, and magnitude of current. The force is found to be independent of contact surface curvature, when contacts that have axial symmetry are considered. Nomographs are given that can ease force calculations for the ideal case, and experimental test results are discussed.

Analysis of Electrodynamic Contact Forces

Two methods are used to derive the equation for the repulsion force: the first is similar to that of Holm,[1] based on a direct calculation of the force by considering lines of current flow in the constriction; the second method is based on energy concepts and uses the principle of virtual displacement to calculate the force.

In the first method, the contact is considered to be a long cylindrical rod with a tapered end, making contact on a flat, infinite plate. The current flow lines are considered as flowing down the axis of the rod, through the tapered section and contact constriction. Fig. 1 is a cross section of this assembly. In the region of the tapered section, the flow lines flow through a large spherical sector of radius $A/\sin\alpha_0$ and terminate on a spherical surface of a small spherical sector of radius $a/\sin\ \alpha_0$.

In Fig. 1 and the relevant equations the following definitions apply:

α_0 = vertex angle of spherical conical sector measured at vertex 0 from the center axis to the tapered contact surface
A = radius of cylindrical rod
a = radius of contact construction at point of contact with infinite plate
θ = angle measured about center axis of rod (not shown in Fig. 1)
r = radial distance of an element of volume $r^2\times\sin\alpha\times dr\times d\alpha\times d\theta$ from vertex 0
α = vertex angle of spherical conical sector of radius r that defines position of element of volume $r^2\times\sin\alpha\times dr\times d\alpha\times d\theta$
I = total current through spherical conical α_0 and contact constriction
I_α = portion of total current through spherical conical α
J_r = current density at any point in spherical conical α at radius r
B_r = flux density existing at an element of volume on edge of spherical α at radius r
μ = permeability of contact medium
ds = surface area of element of volume at radius r perpendicular to radius line r; $ds = r^2\times\sin\alpha\times d\alpha\times d\theta$
df = increment of force acting on element of volume perpendicular to radius r

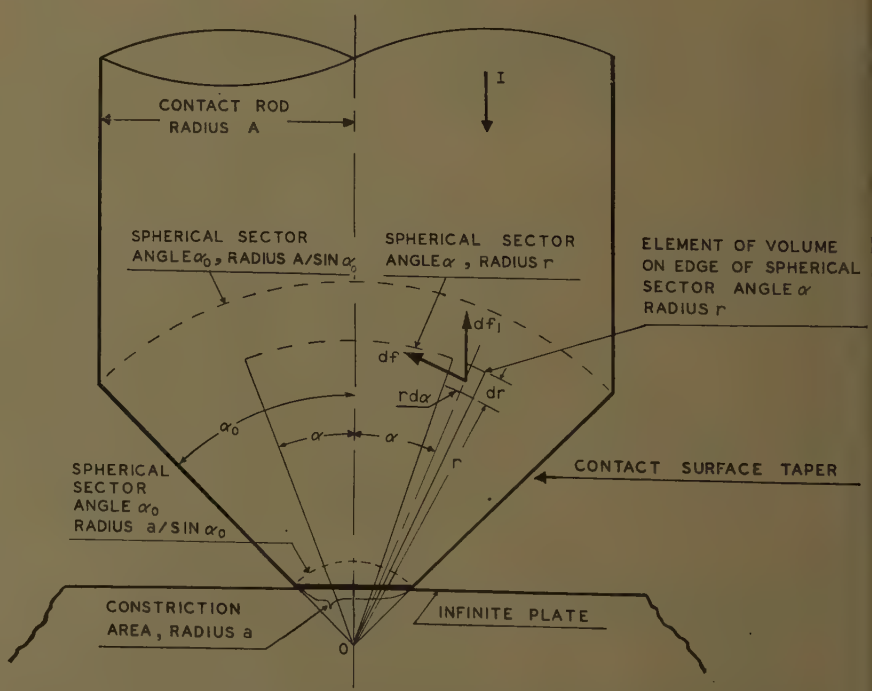


Fig. 1. Cross section of tapered rod making contact on infinite plate

Paper 60-637, recommended by the AIEE Industrial Control Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Great Lakes District Meeting, Milwaukee, Wis., April 27–29, 1960. Manuscript submitted January 21, 1960; made available for printing July 29, 1960.

A. C. Snowdon is with Cutler-Hammer, Inc., Milwaukee, Wis.

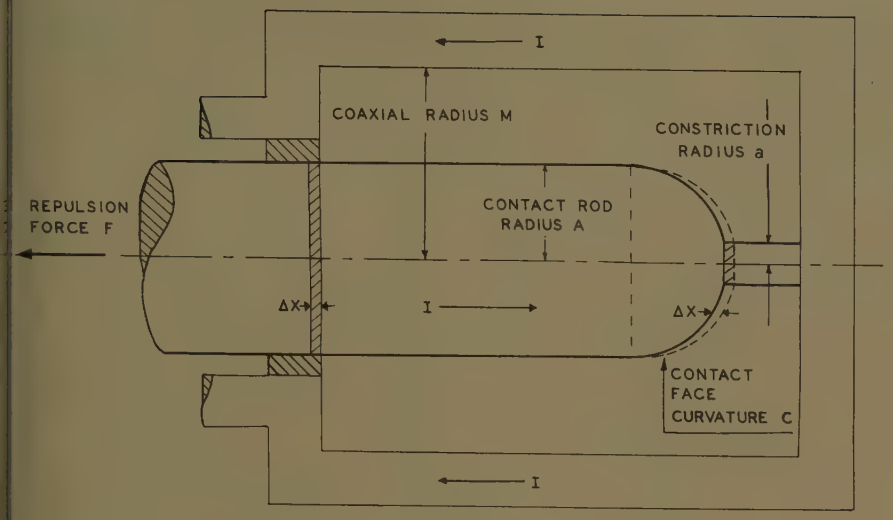


Fig. 2. Coaxial contact circuit element

f_1 = component of df acting parallel to axis of contact rod

f = integral of df_1 over contact volume equal to total electrodynamic repulsion force acting parallel to axis of contact rod

Uniform J_r is assumed at any r from ertex 0 of the spherical sector α_0. A ortion of this current ($J_r \times dS$) will pass hrough the element volume $r^2 \times \sin\alpha \times r \times d\alpha \times d\theta$ and will react electrodynamically with the internal $\mathbf{B}_r$ at this oint. The elemental force df produced is erpendicular to r. A component force $f_1 = \sin\alpha \times df$ exists that reacts parallel o the rod axis. The total contact repulion force $\mathbf{F} = \int df_1 \times dv$ (integral df_1 over olume of spherical α_0) reacts on the conact in a direction parallel to the rod xis.

$\mathbf{B}_r$ can be calculated by means of vector otential and total circuit current disribution. However, if current flow hrough the contact rod is assumed to be ymmetrical with the rod axis, and the eturn path displaced a great distance, Ampere's law can be used by considering he amount of current flowing through the pherical α.

Now, $\mathbf{B}_r$ can be calculated as follows:

$$B_r = \frac{\mu I_\alpha}{2\pi r(\sin\alpha)} = \frac{\mu I}{2\pi r}\,\frac{(1-\cos\alpha)}{(1-\cos\alpha_0)\sin\alpha} \qquad (1)$$

where

I_α = total current × (surface area of conical)/(surface area of conical α_0)

$$I_\alpha = \frac{I(1-\cos\alpha)}{1-\cos\alpha_0} \qquad (2)$$

Now, J_r is total current/(surface area f spherical α_0 at r):

$$J_r = \frac{I}{2\pi r^2(1-\cos\alpha_0)} \qquad (3)$$

Then, to calculate the repulsion force,

$$df_1 = df\sin\alpha = \mathbf{B}_r \times J_r ds \times dr \times \sin\alpha \qquad (4)$$

where

$$df = \mathbf{B}_r \times J_r ds \times dr \qquad (5)$$

Substituting equations 1 and 3 into equation 4 yields

$$df_1 = \frac{\mu I^2}{4\pi^2 r}\,\frac{(1-\cos\alpha)}{(1-\cos\alpha_0)^2}\sin\alpha\, dr\, d\alpha d\theta \qquad (6)$$

$$\mathbf{F} = \int_V df_1 = \frac{\mu I^2}{4\pi^2}\int_{\frac{a}{\sin\alpha_0}}^{\frac{A}{\sin\alpha_0}} \frac{dr}{r} \times \int_0^{\alpha_0} \frac{(1-\cos\alpha)\sin\alpha\, d\alpha}{[1-\cos\alpha_0]^2} \int_0^{2\pi} d\theta \qquad (7)$$

$$= \frac{\mu I^2}{2\pi}\int_{\frac{a}{\sin\alpha_0}}^{\frac{A}{\sin\alpha_0}} \frac{dr}{r} \times \frac{1}{2}\left(\frac{\cos^2\alpha_0 - 2\cos\alpha_0 + 1}{\cos^2\alpha_0 - 2\cos\alpha_0 + 1}\right)$$

$$= \frac{\mu I^2}{4\pi}\ln\left(\frac{A}{\sin\alpha_0} \times \frac{\sin\alpha_0}{a}\right) = \frac{\mu I^2}{4\pi}\ln\frac{A}{a} \qquad (8)$$

The repulsion force is found to depend on I^2, μ of the contact material, and on the natural logarithm of the ratio of the contact radius A and constriction radius a. Since the portion of the integral that concerns the taper angle α_0 is unity, the force is independent of this taper angle.

A derivation of equation 8 based on energy concepts can demonstrate that the repulsion force is independent of an arbitrary contact surface curvature. If the contact and its return circuit are considered as a circuit element, the repulsion force can be calculated by considering energy changes in this circuit element when a virtual displacement of the contact is made.

Fig. 2 is a cross section of the contact and current return path formed into a coaxial circuit element. The return path is shaped as a hollow tube closed at one end. The contact is a movable rod placed in the center of the tube making contact at the closed end. The contact surface curvature of the rod is quite arbitrary, restricted only to having symmetry with the rod axis. The arrows show the direction of current flow in Fig. 2, where

I = total current through contact
M = radius of current coaxial return path
A = radius of contact rod
a = radius of contact constriction, considered as a small cylinder
c = region of contact curvature of an arbitrary shape, but symmetrical with the center axis of the rod

When the contact moves a virtual distance Δx, changes of energy in three areas occur as follows:

1. Change of energy stored in the circuit element magnetic field identified as W_i.
2. Work done by contact in moving virtual distance Δx identified as W_m.
3. Change of energy supplied by the external circuit identified as W_e.

Then, by the principle of energy conservation,

$$W_e - W_m = W_i \qquad (9)$$

The contact circuit element is driven by a current source, since the impedance of a contact is very low compared with that of the rest of the circuit. This means that changes in contact impedance cause little change in current; hence the current through the contact can be assumed constant during the virtual movement Δx.

If the virtual movement Δx occurs in time ΔT, the electric energy into the contact element during this time is

$$W_e = \int_0^{\Delta t} I\,E\,dt \qquad (10)$$

where

E = induced voltage across contact element caused by changing flux linkages when virtual movement is made
I = instantaneous line current through contact assumed to be constant over time Δt

Then,

$$W_E = \int_0^{\Delta T} I\,d\frac{(LI)}{(dt)}dt = \int_{L_1}^{L_2} I^2 dL + \int_{I_1}^{I_2} LI dI$$

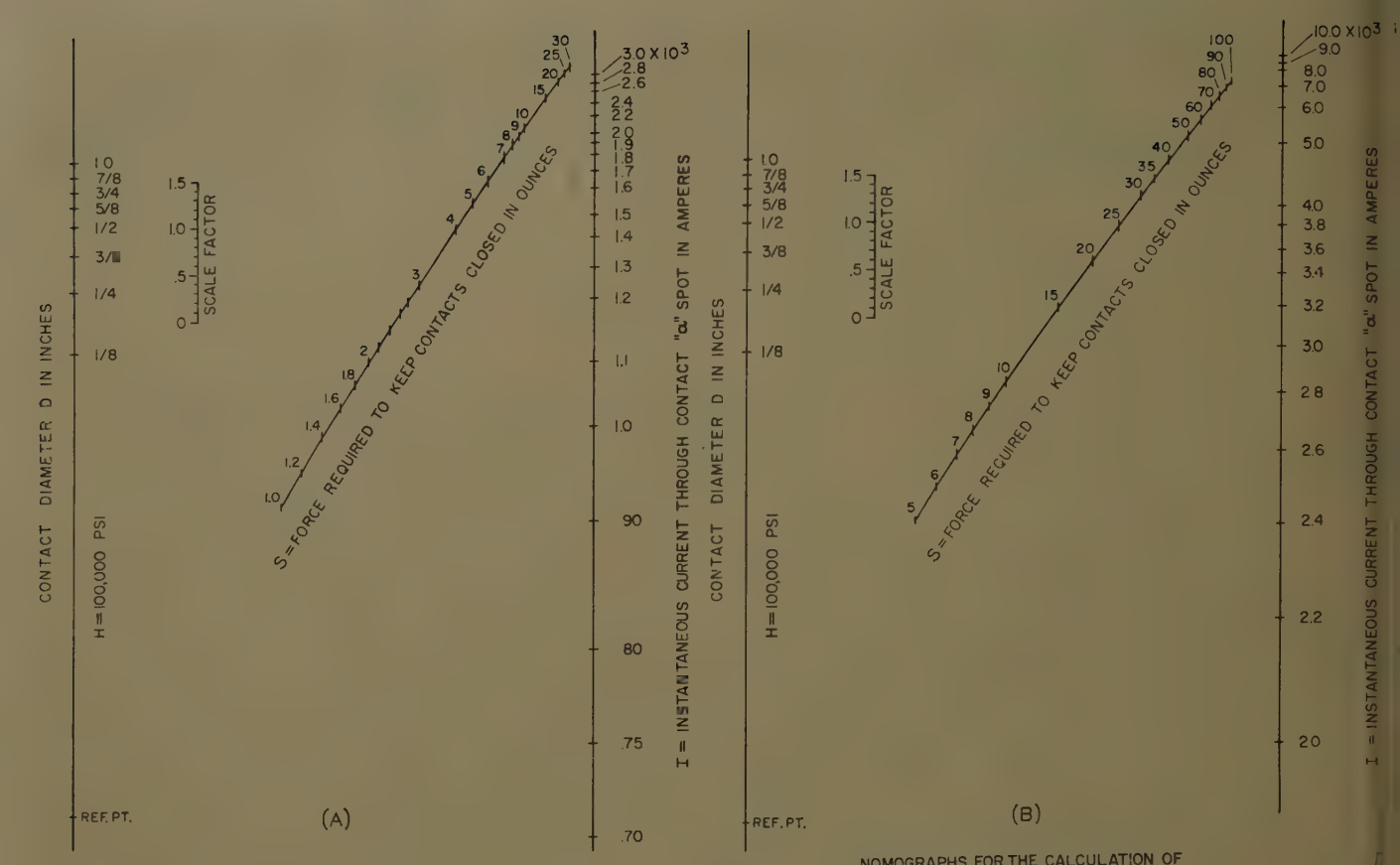
Since $I_1 = I_2 =$ a constant,

$$W_E = I^2(L_2 - L_1) = I(\lambda_2 - \lambda_1) \qquad (11)$$

where

L_1, λ_1 = inductance or flux linkages of the contact circuit element at zero reference time
L_2, λ_2 = inductance or flux linkages of the contact circuit element at time ΔT

The change of internal energy W_i of the circuit element is proportional to the difference in its inductance before and after the virtual contact movement.



NOMOGRAPHS FOR THE CALCULATION OF CONTACT ELECTRODYNAMIC REPULSION FORCES

$$F = 36 \times 10^{-8} I^2 \ln\left(2\sqrt{\pi H} \times \frac{D}{\sqrt{P}}\right)$$

$$\text{SCALE FACTOR} = \frac{1}{2}\ln\left(\frac{H}{100{,}000}\right)$$

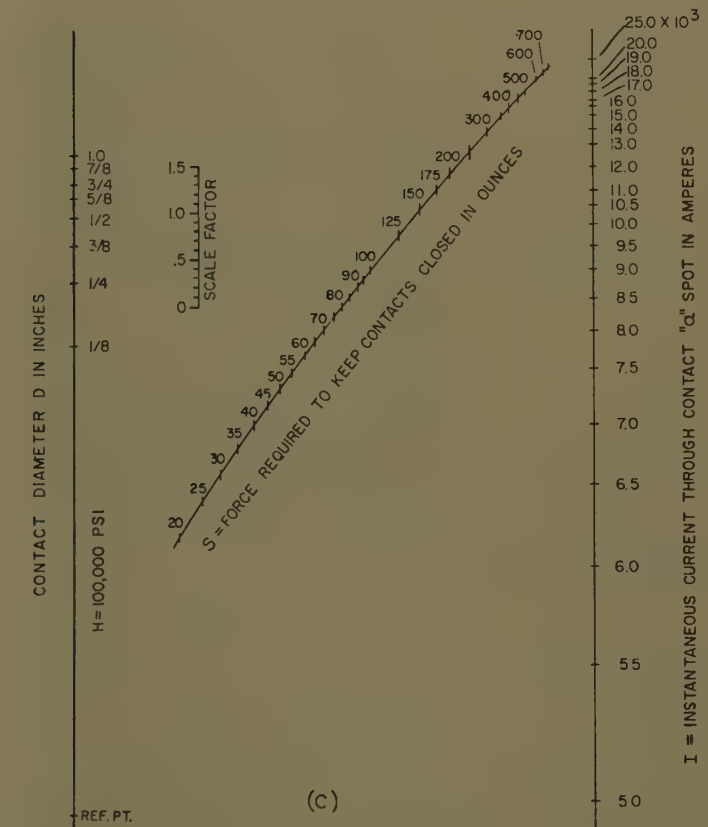


Fig. 3. Repulsion force nomographs for currents, in amperes, up to: A—3,000; B—10,000; C—25,000

$$W_1 = \frac{1}{2} I^2 L_2 - \frac{1}{2} I^2 L_1 = \frac{1}{2} I^2 (L_2 - L_1)$$

$$= \frac{1}{2} I(\lambda_2 - \lambda_1) \quad (12)$$

The mechanical energy out of the circuit element caused by virtual contact displacement during time ΔT is

$$W_m = \int_0^{\Delta x} \mathbf{F}\, dx = \mathbf{F}\Delta x \quad (13)$$

where

$\mathbf{F}$ = repulsion force

Δx = virtual movement of contact during time T

Substituting equations 11, 12, and 13 into equation 9 yields

$$I[\lambda_2 - \lambda_1] - \frac{1}{2} I(\lambda_2 - \lambda_1) = \frac{1}{2} I(\lambda_2 - \lambda_1) = F\Delta x \quad (14)$$

The problem reduces to one of calculating the flux linkages of the contact coaxial circuit element before and after a virtual movement Δx. Since the system is coaxial, the simple equation commonly applied to calculate total flux linkages per unit length of a coaxial wire can be used.[3]

$$\frac{\lambda}{\text{length}} = \frac{\mu I}{2\pi}\left[\left(\ln \frac{M}{r}\right) + \frac{1}{4}\right] \quad (15)$$

where M is the inside radius of the coaxial return path and r the radius of the inner conductor. The term $\mu I/2\pi \times 1/4$ defines the internal flux linkage of the inner conductor. This internal flux linkage is independent of conductor radius if a uniform current density over a cross section is assumed.[3]

When the contact of Fig. 2 is moved a distance Δx, the flux linkages are changed in two areas: the portion of flux linkages

rrounding contact rod radius are de-eased by an amount $\mu I \Delta x/2\pi[(\ln M/-[+1/4]$. The flux linkages of the lindrical conductor comprising the ntact constriction are increased by an nount

$$\frac{\Delta x}{\pi}\left[\ln\left(\frac{M}{A}\right)+\frac{1}{4}\right]$$

he flux linkages of region C, defining the bitrary face curvature of a symmetrical ntact, remain unchanged during a ovement of Δx.

The net change in flux linkages with is movement are

$$\frac{\Delta x}{2\pi}\left[\left(\ln\frac{M}{a}\right)+\frac{1}{4}-\left(\ln\frac{M}{A}\right)-\frac{1}{4}\right] \qquad (16)$$

$$\frac{\mu I \Delta x}{2\pi}\left(\ln\frac{A}{M}\times\frac{M}{a}\right)=\frac{\mu I \Delta x}{2\pi}\ln\frac{A}{a}$$

hen

$$I(\lambda_2-\lambda_1)=\frac{\mu I^2 \Delta x}{4\pi}\ln\frac{A}{a}=F\Delta x \qquad (17)$$

$$=\frac{\mu I^2}{4\pi}\ln\frac{A}{a}$$

quation 17 is the same as equation 8 de-ved by the first method.

ractical Calculations

Equations 8 and 17 demonstrate that he repulsion force is dependent on the ntact constr ction radius a. Holm[1] hows that the magnitude of a can be pproximated by considering that the ressure between contacts causes the ridging constriction to be in a state of lastic deformation that supports the oad. It is assumed that the contacts ouch in only one point and that the area f contact is circular. The equation re-ating this contact area to a material con-tant is given by Holm[4] and Shobert[5] to e

$$P_1=H_1 A_a \qquad (18)$$

here

A_a = constriction area, inches^2
H_1 = contact material hardness, pounds per inch^2 (see references 4 and 5)
P_1 = contact pressure in pounds

f the area A_a is assumed to be circular, hen the radius a of the constriction is

$$a=\sqrt{\frac{P_1}{\pi H_1}} \qquad (19)$$

The nomograph calculation methods hat follow use a constriction radius cal-culated from equation 19 and are based on equations 8 or 17. Use of these equations

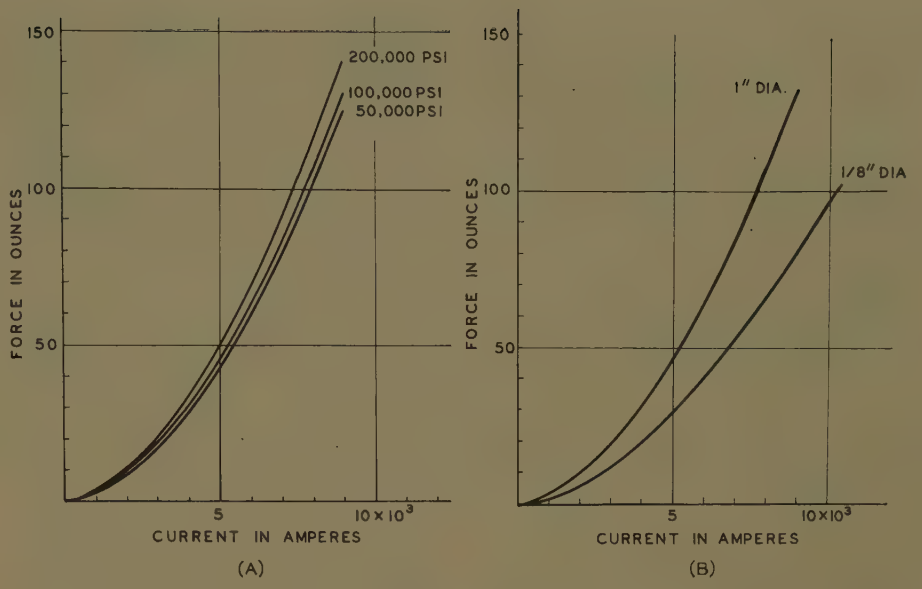


Fig. 4. Calculated comparison of influences on repulsion force as functions of current. A—Influence of H, for contacts of 1-inch diameter. B—Influence of contact diameter, H = 100,000 psi

requires that the magnitude of constriction radius be determined by the initial contact pressure, and that the radius remains at this magnitude until the contacts separate or "blow apart." Actually, this idealized condition is not realized in practice. However, experiment shows that the idealized momograph calculations can be modified to give meaningful results.

If practical measurement units and the magnitude of a from equation 19 are substituted into equation 8 or equation 17, the equation from which the nomographs in Figs. 3(A), (B), and (C) were designed is obtained. This equation is

$$\mathbf{F}=S=36\times10^{-8}I^2\ln\left(\frac{2\sqrt{\pi H}}{1}\times\frac{D}{\sqrt{P}}\right) \qquad (20)$$

where

F is expressed in ounces
S = applied contact pressure ounces
P = net contact pressure equal to difference between S and **F**, ounces; $P=S-\mathbf{F}$
H is expressed in psi (pounds per square inch)[1,5]
D = contact diameter, inches
I = instantaneous current through contact constriction or a spot, amperes

The nomographs were designed on the basis that the contacts separate when $\mathbf{F}=S$.

Use of Nomographs

The nomographs given in Figs. 3 (A), (B), and (C) are designed for various contact diameters of materials with an H of 100,000 psi. The force is determined by laying a straight edge from the value of contact diameter to the value of instantaneous current through the contact constriction or a spot. The force required to keep the contacts from separating is read on the centers cale.

It should be noted there is a separate calibration called a scale factor, which is used for making calculations involving materials with an H different from 100,000 psi. The determination of H for materials can be made from superficial Rockwell hardness readings as shown by Shobert.[5] The scale factor (SF) is calculated by

$$\mathrm{SF}=1/2\ln\left(\frac{H}{100{,}000}\right) \qquad (21)$$

This magnitude is located on the SF scale on the appropriate nomograph. The dimension is transferred with dividers to the D scale. The SF dimension is added or subtracted to or from the location of the required D. This new location is used to make calculations for the required D.

As an example, suppose it is desired to determine the force on a copper contact with an H of 175,000 psi, 1-inch diameter, carrying an instantaneous current of 5,000 amperes. The SF calculated by equation 21 is

$$\mathrm{SF}=\frac{1}{2}\ln\frac{(175{,}000)}{(100{,}000)}=+0.278 \qquad (22)$$

With the use of a pair of dividers the SF dimension +0.278 is added to the 1-inch location on the D scale, which is the point to use to calculate the contact repulsion force for a material of an H=175,000 psi, one inch in diameter. A line is drawn from this point to the instantaneous current of 5,000 amperes, after which 50 ounces is read on the force scale.

Fig. 4(A) shows a family of curves giving calculated contact repulsion force as a function of current for various values of H. The curve shows that a variation of

Table I. Experimental Results

Current, Amperes	Force,* Ounces Calculated	Measured	Per Cent Difference
5,500	65	80	19
6,000	70	72	2
5,100	52	66	21
5,700	65	66	2
5,600	61	64	5
4,800	44	45	3
4,200	36	45	20
5,000	49	61	20

* Contact pressure required to keep contacts from separating when carrying given current.

H of 4 to 1 causes a change in repulsion force on the order of 20%.

Fig. 4(B) shows calculated repulsion force as a function of current for contact diameters of 1/8 inch and 1 inch of a material with an H of 100,000 psi. Examination of the curves shows that this variation in diameter causes a repulsion force change on the order of 60%.

Experimental Results

Table I lists the results of tests performed on 1-inch-diameter copper contacts. It is evident that the contact pressure needed to hold the contacts closed varies by about 20%. Part of this variation can be attributed to errors in measurement of constriction radius a. If contact conditions are ideal, so that the constriction area is circular and in a state of plastic deformation, measurement of contact pressure should give a reasonably accurate value of constriction radius. However, as Holm[4] shows, the constriction region may be partially in a state of elastic deformation as well as being plastically deformed. This deviation from the ideal case in a practical contact causes errors in calculating the constriction radius a by equation 19. This error would be such that a smaller value of repulsion force is calculated by the nomographs than actually exist.

The magnitude of contact mass and pressure-spring rate can also cause a variation of measure repulsion force from that calculated by the nomographs. When the repulsion force is exerted on the contact, the contact will start to move in accordance with the dynamic properties of contact mass and pressure-spring constant. The movement will reduce the size of the constriction radius a, which will cause a greater repulsion force at a given current than would normally exist. This increased force will cause an increase in contact acceleration and movement, leading again to an increase in repulsion force. The net result will be to part the contacts at a lower value of current than is calculated by the nomographs.

Even though these factors cause force magnitudes calculated by the nomographs to disagree with experimentally measured values, the nomographs provide a good tool to estimate contact pressure force requirements. Experience with a variety of contact structures demonstrates that for any given class of contact structure a multiplying correction factor can be applied to the results. For example, a multiplying factor of 1.2 applied to the calculated data shown in Table I would insure that the calculated contact pressure determined would keep the contacts closed.

Conclusions

The analysis has provided theoretical insight into the nature of the contact repulsion force to the extent that this force is independent of contact face curvature with contacts shaped symmetrically about the axis. In addition, nomograph calculation aids have been provided that estimate repulsion force magnitude. However, more study is needed to account for the variance between measured and calculated force magnitudes. The foremost problem is that of analyzing in detail the effect of contact mass and pressure-spring constant on the repulsion force.

References

1. Electric Contacts Handbook (book), R. Holm. Springer-Verlag, Berlin, Germany, 1958, pp. 55–56.

2. Calculation of Contact Electrodynamic Forces Caused by Current Constriction and Associated Geometries, A. C. Snowdon. *Engineering Proceedings P–36*, "Electrical Contacts—1959," Pennsylvania State University, University Park, Pa., June 1959, pp. 70-84.

3. The Electro-Magnetic Field in Its Engineering Aspects (book), G. W. Carter. Longmans, Green and Company, London, England, 1954, pp. 215–17.

4. See reference 1, pp. 35–38.

5. Calculations of Electrical Contacts Under Ideal Conditions, Erle I. Shobert, II. *Proceedings*, American Society for Testing Materials, Philadelphia, Pa., vol. 46, 1946, pp. 1126–46.

Heat Transfer Aspects of Electric Heating

B. ROLSMA
NONMEMBER AIEE

THIS PAPER is a discussion of the principles of steady-state heat transfer as they apply to the field of electric heating and, as such, no solutions are derived, but rather the governing differential equations are presented and references containing the solutions for various boundary conditions are cited. In most of the references given, the work was concerned with the temperature distribution in electric coils or other electric apparatus wherein the generation of heat was an unwanted phenomenon and not, as is considered here, the prime function of the device. However, since these are the solutions to basic differential equations, they are directly applicable to any other body for which the basic equation and the boundary conditions hold.

An attempt has been made to divide the discussion into essentially two parts; the first part deals with the heat transfer within electric heating elements themselves, and the second part is devoted to heat transfer from the surfaces of the elements. Since the temperature distribution within an electric heating element depends upon the heat transfer rate from its surfaces, the two are not physically separable, but, in practice, an engineer applying purchased electric heating elements is interested in dissipating the developed energy from the surfaces of the elements and not in the internal heat transfer problems.

Exact Solutions for Heat Transfer in Electric Heating

The differential equations for the steady-state temperature distribution in an electric heating element are derived by performing a heat balance upon a

Paper 61-30, recommended by the AIEE Electric Heating Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 29–February 3, 1961. Manuscript submitted October 31, 1960; made available for printing December 28, 1960.

B. Rolsma is with the General Electric Company, Shelbyville, Ind.

The author is indebted to Professor S. P. Kezios, Illinois Institute of Technology, and consultant to the Industrial Heating Department, General Electric Company, for making available copies of several of the references cited; also to the General Electric Company for all the photographs used.

fferential volume in the solid material. ssuming that the medium is homo- neous and neglecting any changes in s thermal conductivity due to tempera- re variations within the heating element, e heat balance on the differential volume elds

$$\left(\frac{\partial^2 t}{\partial x^2}+\frac{\partial^2 t}{\partial y^2}+\frac{\partial^2 t}{\partial z^2}\right)+q'''=0 \qquad (1)$$

If the internal heat generation per nit volume, the q''' term, is considered be constant, this equation is exactly oisson's equation.

Solutions to equation 1 for simple geo- etric configurations with fixed surface mperatures and constant q''' are well nown. Rogowski[1] has shown that the mperature distribution within infinitely ide plates, infinitely long cylinders, and pheres has a parabolic form. Jakob[2] reated the infinitely long rectangular bar ith these same boundary conditions. he results of his work expressed as a atio of the temperature distribution in n infinitely wide plate are summarized n reference 3.

The assumption of a negligible change n the thermal conductivity with tempera- ure as used in the derivation of equation 1 oupled with the use of a constant q''' erm in producing the solutions cited is uite realistic for electric resistance eating element materials. The Wiede- nann-Franz-Lorenz equation indicates hat thermal conductivity and electric onductivity are related by a simple xpression of temperature multiplied by orenz number. Since the Lorenz num- er does not exhibit large changes with emperature variations, thermal conduc- ivity and electric conductivity are essen- ially related by an expression of a con- tant multiplied by temperature and, ence, the ratio, q'''/k, in equation 1 may e expected to follow this same relation- hip. The problem then is to select a alue of q'''/k in the range of tempera- ure encountered in the problem.

Although it is recognized that both hermal conductivity and heat generation n electric resistance heating elements vary with temperature, solutions to equation 1 ave been developed considering that the $'''$ term alone varies with temperature. akob[4] generated solutions to equation 1 or the infinitely wide plate, the infinitely ong cylinder, and the sphere with fixed urface temperatures and a linear varia- ion of q''' with temperature, while Higgins[5] investigated the temperature distribution in an infinitely long rec- angular bar under these same condi- ions.

When the changes in thermal conductivity with temperature variation cannot be negelected in a homogeneous material, the steady-state heat balance on the dlfferential volume yields the relationship

$$k\left[\frac{\partial^2 t}{\partial x^2}+\frac{\partial^2 t}{\partial y^2}+\frac{\partial^2 t}{\partial z^2}\right]+q'''+\frac{\partial k}{\partial t}\left[\left(\frac{\partial t}{\partial x}\right)^2+\left(\frac{\partial t}{\partial y}\right)^2+\left(\frac{\partial t}{\partial z}\right)^2\right]=0 \qquad (2)$$

This equation is clearly nonlinear. However, the use of the transform procedure introduced by Kirchhoff[6] in 1894 will reduce equation 2 to a linear equation similar in form to equation 1. Applications of this procedure are to be found in the discussion by Jakob of Higgins' paper[7] and in reference 8.

The transform method is applicable to the cases of varying thermal conductivity with a constant q''' term or with a temperature-dependent q''' term. When it is considered that both thermal conductivity and q''' vary with temperature, both k and q''' are expressed in terms of the transform variable and substituted in equation 2. This substitution will reduce equation 2 to Poisson's equation. On the other hand, when thermal conductivity alone varies, the solutions obtained for equation 1 with a constant q''' term may be directly applied to equation 2 by introducing the transform variable into these solutions provided the boundary conditions are those of fixed surface temperature or thermal conductivity multiplied by the temperature gradient at the surface.

The heat balance for the more general case of heat transfer with thermal conductivity varying with both temperature and direction in the differential volume yields

$$\frac{\partial}{\partial x}\left(k_x\frac{\partial t}{\partial x}\right)+\frac{\partial}{\partial y}\left(k_y\frac{\partial t}{\partial y}\right)+\frac{\partial}{\partial z}\left(k_z\frac{\partial t}{\partial z}\right)+q'''=0 \qquad (3)$$

Reference 8 contains a solution for this case considering that the q''' term varies linearly with temperature and, further, that the heat transfer rates at the four surfaces of the rectangular bar examined vary, i.e., both the surface heat transfer coefficients and the sink temperatures are different for the four surfaces.

The exact analytical solution of the general case of heat transfer described by equation 3 with variable heat generation and the boundary conditions of heat transfer with different surface coefficients and sink temperatures on all faces of the body for even the more simple geometric shapes is a very difficult and time-consuming task. Quite often it is not even possible to obtain the solution by exact analytical means and finite difference techniques must be used. However, before leaving the area of exact analytical solutions, the work of Carslaw and Jaeger[9] must be mentioned for it is the most complete compilation of mathematically rigorous solutions to the problems being considered here.

Approximate Solutions for Heat Transfer in Electric Heating

In the numerical analysis of heat transfer problems, the differential equations 1, 2, and 3 are replaced by difference equations of the general form

$$\left[k_x\frac{t_{(x+\Delta x,y,z)}-2t_{(x,y,z)}+t_{(x-\Delta x,y,z)}}{(\Delta x)^2}\right]+k_y\left[\frac{t_{(x,y+\Delta y,z)}-2t_{(x,y,z)}+t_{(x,y-\Delta y,z)}}{(\Delta y)^2}\right]+k_z\left[\frac{t_{(x,y,z+\Delta z)}-2t_{(x,y,z)}+t_{(x,y,z-\Delta z)}}{(\Delta z)^2}\right]+q'''_{(x,y,z)}=0 \qquad (4)$$

Schneider[10] has applied numerical techniques to yield solutions for the temperature fields in electrical coils, while Dusinberre[11] presents a very complete case for the solution of heat transfer problems in general by numerical analysis.

While the relaxation method originated by Southwell[12] and initially applied to problems in the field of heat transfer by Emmons[13] is somewhat tedious, it will produce acceptable results provided the body subdivisions, Δx, Δy, and Δz, are small. As Δx, Δy, and Δz become extremely small, the results of the relaxation process approaches the exact solution for a given problem, and, in the limit, as Δx, Δy, and Δz approach zero, equation 4 becomes equation 3 exactly. However, as Δx, Δy, and Δz become smaller in a given region, more and more finite difference equations are introduced. Some compromise between the desired accuracy of results and the time or labor involved in obtaining a solution is then in order.

When a given problem is first encountered, there is no sure way of judging the errors introduced by selecting a certain body subdivision. The only course open is to select one subdivision size and to obtain an initial solution. The analysis is then repeated with a subdivision size, say one-half the previous one. A wide difference in the results of these two solutions would indicate that a further reduction in subdivision size is required; while a very small difference in results indicates that a larger subdivision could possibly be used. It is only through experience of this nature that a basis for selecting subdivision size can be derived.

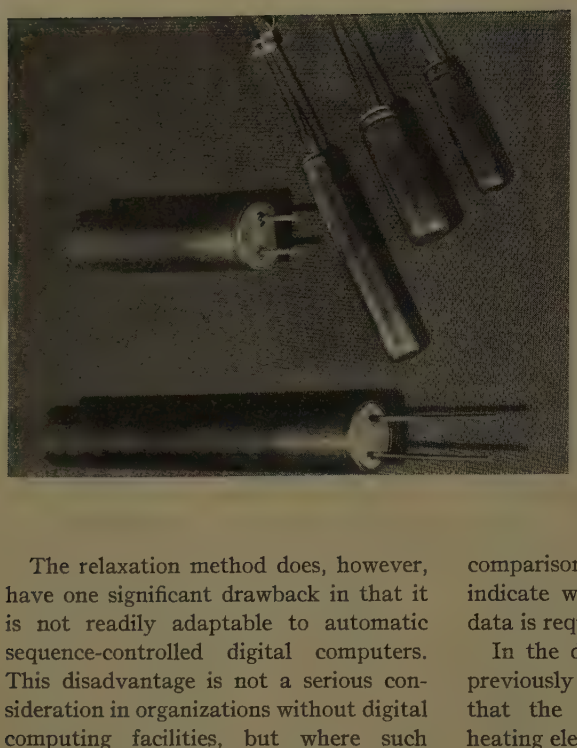


Fig. 1 (left). Various sizes of General Electric Company cartridge - type heaters

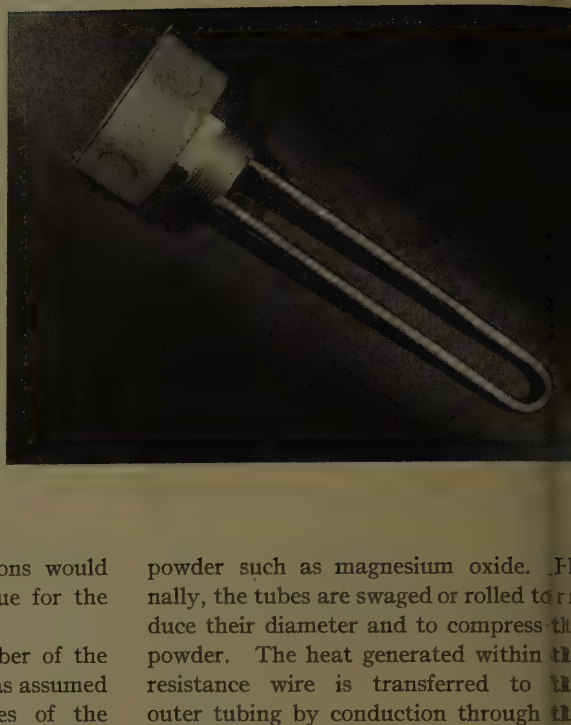


Fig. 2 (right). 1,000-watt water immersion heater

The relaxation method does, however, have one significant drawback in that it is not readily adaptable to automatic sequence-controlled digital computers. This disadvantage is not a serious consideration in organizations without digital computing facilities, but where such equipment is available, other methods are being employed to solve the finite difference equations.

One of the methods frequently used in conjunction with digital computers to solve finite difference heat flow equations is the iteration method.[14] Although the iteration process, when performed manually, is generally much slower than the relaxation process, general programs may be written to direct a digital computer to automatically and rapidly perform the calculations required by the iteration process and free the analyst of this task.

The speed of the digital computer and its ability to store and process much data somewhat eliminates the need for the compromise between desired accuracy of results and subdivision size mentioned before. Using the digital computer, it is possible to evaluate quickly the effect of subdivision size on the results and to actually "home in" on a solution by repeatedly reducing the size of the subdivisions and rerunning the problem through the computer. The only apparent restriction would be computer memory size and the time required to complete the data input forms necessary for the program.

The digital computer is then a powerful tool, but a word of caution: The results of the calculations can be no better than the data used. Again, in the case of uncertain data, the speed of the digital computer does offer an advantage in that a problem may be rapidly solved with first the high value of the uncertain data and then with the low value of the data. A comparison of the two solutions would indicate whether a better value for the data is required.

In the derivation of a number of the previously cited solutions, it was assumed that the surface temperatures of the heating elements were fixed and invariant over the individual surfaces. This may not always be the case in practice, for the surfaces of a heating element and even discrete portions of these surfaces will assume whatever temperature may be required to dissipate the heat generated. If the heat generated within the element cannot be dissipated at temperatures below the melting point of the material then the element will fail. Thus, an element designed for a certain ambient condition and installed under a different condition may fail very quickly in service if the installation conditions have an adverse effect on the dissipation of heat from the surfaces of the element. It is imperative, then, that the design engineer be informed of all the conditions of installation when custom-designing heating elements, and that the equipment design engineer, when applying "off-the-shelf" heating elements, be fully aware of the design conditions of these heating elements.

Heat Transfer from Electric Heating Elements

All of the various modes of heat transfer, conduction, free and forced convection, and radiation, may be active to one extent or another in removing heat from the surfaces of electric heating elements and in some cases all may be important. As an example of this, consider the tubular-type heating element found in the modern electric range. This unit is fabricated by centering a helically wound section of resistance wire within the element sheath and then filling the space remaining inside the tube with an electrical insulating powder such as magnesium oxide. Finally, the tubes are swaged or rolled to reduce their diameter and to compress the powder. The heat generated within the resistance wire is transferred to the outer tubing by conduction through the insulating powder and, in addition, by radiation between particles. Heat is dissipated, in turn, from the outer tubing or sheath, by radiation and free convection.

In some cases, conduction is the principal heat transfer mechanism removing heat from the outer sheath of heating elements. Low-temperature, 750 F (degrees Fahrenheit) maximum, cartridge heaters of the type shown in Fig. 1 are inserted in holes in die blocks, press platens, molds, and other bodies which must be heated. These heaters are designed to fit snugly in a hole of a specified size and to dissipate their energy through the gas film in the clearance space between the heater and the sides of the hole to the body being heated. Since radiation is not very effective at these low temperatures, almost all of the heat will be transferred by conduction through the gas film. If by some chance the size of the hole is increased slightly so that the clearance is twice the design value, the temperature differential between the heater sheath and the surfaces of the hole will be approximately twice the design value. This increase in temperature differential may cause the temperature within the heater to rise beyond the maximum permissible.

Immersion-type heaters, Fig. 2, are designed to dissipate their heat by natural convection to liquids and, hence any installation condition which inhibits the free circulation of the liquids over the heater element surfaces is to be avoided. In regard to this, locations which place immersion-type heaters in close proximity to the sides of the containing vessel

ear the free surface of the liquid being eated are poor.

It should also be recalled that the liquid hysical properties—viscosity, thermal nductivity, specific heat, and density—re important factors in determining the rface coefficient of heat transfer for both rced and free convection. Thus, an mmersion heater designed for heating ater to a certain temperature may not e suitable for heating other liquids to this me temperature, since 4-to-1 variations the natural convection coefficient beween different liquids is not uncomon.

Attention must be directed to the rinciples of forced-convection heat transr and fluid flow when heaters of the types own in Figs. 3 and 4 are used to heat ases flowing in ducts. When heaters re to be installed in ducts or boxes rough which gases flow, configurations milar to those encountered in the degn of heat exchangers are involved. Much information on the flow characterstics and heat transfer rates from bare nd finned tubes is to be found in the echnical literature.

The flow information obtained from ests of heat exchanger configurations an be directly applied whenever these me configurations occur in the applicaion of electric heating elements. Thus, he bypass effects found in heat exchangrs will also be present in electric heating pplications. The spacings required to btain good mixing between rows of eat exchanger tubes will also be required etween rows of electric heating elements nd the pressure drop through rows of lectric heating elements may be calulated from correlations for the pressure rop in heat exchangers.

On the other hand, the heat transfer data derived from heat exchanger studies cannot be directly used for determining electric heating element designs. Schmidt and Wenner[15,16] have shown that the local surface heat transfer coefficients vary along the circumference of a cylinder when flow is normal to its axis. This inherent variation in the local coefficient around the surface of a cylinder can readily, and indeed does, create localized hot spots on the surface of cylindrical heating elements in cross-flow. This variation in coefficient is neglected in correlating data from heat exchanger investigations and the values of heat transfer coefficient reported are average values taken over the entire surface of the tubes. These coefficients are then not suitable for electric heating element design, for it is the coefficient based on hot-spot temperature which must be used to determine the electric heater rating.

One additional consideration, not present in the design of heat exchangers, does exist in selecting and arranging electric heating elements for heating gases, or for that matter, any fluid flowing in a duct or pipe. When large temperature rises are required in the fluid flowing through the heater bank, it may be necessary to derate or reduce the total power input to the heaters in the last few rows of the heater bank so that their temperatures do not exceed the maximum allowable.

When temperatures in excess of 700 F are encountered in electric heating, heat transfer by radiation begins to play an important role. While it is well recognized that the temperature level of operation is a prime factor in establishing the rate of heat transfer by radiation, there are other factors which must also be considered. The geometry of the heating element, the geometry of the heat receiver, and their positions relative to each other and to any other surfaces viewed by either the source or the receiver are factors which also influence the rate of radiation heat transfer. In addition to these geometric considerations, the surface characteristic, emissivity, must be examined for not only the energy source and the ultimate receiver of this energy, but for all surfaces in the radiation system.

Much information is available concerning the geometry factors in radiation heat transfer.[17–19] Emissivity values for both unoxidized and oxidized metallic surfaces and for nonmetallic surfaces are to be found in references 20 and 21.

The solution to problems involving radiation heat transfer can become quite complex when more than three surfaces are involved in the interchange. Not only will the consideration of additional surfaces make the problem more complex, but the simplest 2-surface problem becomes complex when variation in the temperatures of these surfaces makes it necessary to subdivide them into zones of uniform, or substantially uniform, temperature. Oppenheim[22] presents a network analysis method for solving the diffuse radiation problem within an enclosure formed by grey surfaces. Hottel[20] presents a slightly different approach to this same problem. These methods, for the steady-state case, result in a series of linear simultaneous equations which can be solved either with a digital computer, if one is available, or with a desk calculator, for up to eight equations,

Fig. 3 (left). Edge-wound resistor heating unit for use in forced-convection heating

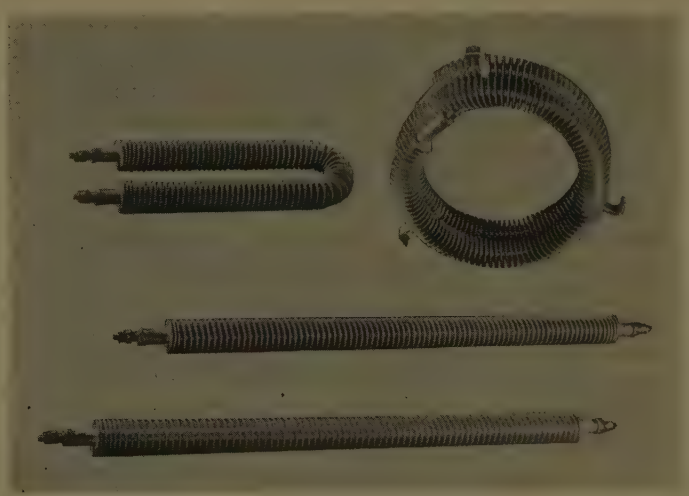


Fig. 4 (above). Finned tubular heaters

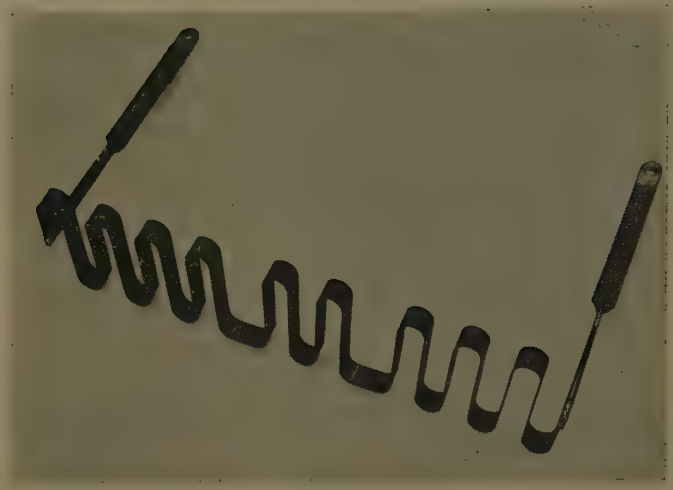


Fig. 5. Electric resistor heating element for furnace service

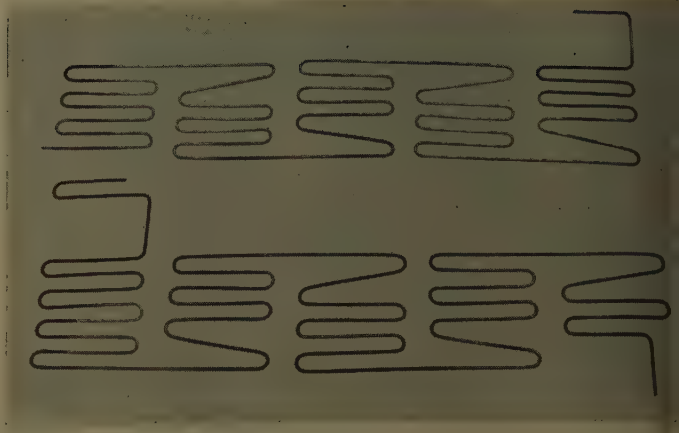


Fig. 6. Molybdenum resistor heating element for high-temperature furnaces

using Crouts[23] method as suggested by Hottel.

Heating elements for high-temperature service in ovens and furnaces are made from both metallic and nonmetallic electric resistance materials. Nonmetallic elements are primarily straight cylindrical shapes with low-resistance sections at the ends where the elements pass through the furnace or oven walls, while metallic heating elements are frequently cast or formed in sinuous shapes. Figs. 5 and 6 show metallic elements formed using resistance alloy ribbon and molybdenum rod respectively. Also shown in Fig. 5 are the low-resistance terminal ends to be passed through the furnace wall.

When designing the metallic resistors, the spacing between adjacent loops must be made large enough to permit the heat generated within the material to radiate freely out into the furnace. If the spacing is too small, the ribbon temperature required to dissipate the energy developed will become unreasonable. The length of each loop is controlled by the ability of the material to support itself at high temperature. If the loops are too long, sagging or excessive elongation of the elements will result.

Attention to the details of supporting the resistance heating elements in the furnace may reduce the operating temperature of the element considerably. When ceramic supports are employed, localized hot spots may develop because of the radiation screening or shading effect of the support and the inability of the ceramic material to conduct heat away from these local hot spots. Metallic hooks are preferred for supporting the elements, since they are usually much smaller than ceramic supports and, hence, do not screen the view from the element to the furnace as much, and have, in addition, the ability to conduct the heat away from a hot spot more readily then ceramic materials.

With adequate support, the only limitation on the temperatures attained by electric heating elements is the melting point of the element material. However, since the oxidation rates at these elevated temperatures would normally result in a very short element life, protective or inert atmospheres are introduced into the furnace, or the furnaces are designed for operation with a vacuum atmosphere. At first glance, the requirement of a protective or inert atmosphere for reasonable element life would appear to be a deterrent to the use of electric heat, but it must be remembered that many of the processes which require the higher temperature levels also require that the processing be done in a protective atmosphere. Actually then, electric heat has an advantage over other methods of heating in this case, since electric elements can operate in protective atmospheres without contaminating the atmosphere, while other methods of heating require the use of an envelope or muffle to separate the heat source from the protective atmosphere and the material being processed. Further, whenever a muffle or envelope is placed between the source of heat and the eventual receiver of heat, the net rate of radiant interchange at a specified temperature level is reduced.

Conclusions

In summary, the heat transfer problems of the electric apparatus designer and the electric heating element designer are closely related, for both are concerned with the same basic differential equations within the heat generating regions of their respective equipments and with removing this heat to the equipment surroundings with as small a temperature differential as possible. The only significant difference between the two is that the electric apparatus designer considers the Joulean heat as an unwanted quantity and seeks to reduce it to an absolute minimum, while the electric heating element designer is constantly striving to increase the heat generated to the maximum possible without sacrificing element life.

Nomenclature

k = thermal conductivity; k_x, k_y, k_z = thermal conductivity in x, y, z direction

q''' = internal heat generation per unit volume

t = temperature

x, y, z = co-ordinate axis directions

Δx, Δy, Δz = increments in the x, y, z directions

References

1. Erganzung der Erwärmungsvorschriften: Zum Vorschlag des Herrn Vidmar, W. Rogowski. *Archiv für Elektrotechnik*, Berlin-Charlottenburg, Germany, vol. 7, 1918–19, pp. 41–47.

2. Die Temperatverteilung in einer elektrischen Wicklung von rechteckigem Querschnitt, M. Jakob. *Ibid.*, vol. 8, 1919–1920, pp. 117–26.

3. Heat Transfer, Vol. 1 (book), M. Jakob. John Wiley & Sons, Inc., New York, N. Y., 1959, p. 179.

4. Influence of Nonuniform Development of Heat Upon the Temperature Distribution in Electrical Coils and Similar Heat Sources of Simple Form, M. Jakob. *ASME Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 65, 1943, pp. 593–605.

5. Formulas for Calculating the Temperature Distribution in Electrical Coils of General Rectangular Cross Section, T. J. Higgins. *Ibid.*, vol. 66, 1944, pp. 665–70.

6. Vorlesungen Über die Theorie der Waerme (book), W. Kirchhoff. 1894, p. 13.

7. Formulas for Calculating Temperature Distribution in Transformer Cores and Other Electrical Apparatus of Rectangular Cross Section, T. J. Higgins. *AIEE Transactions*, vol. 64, Apr. 1945, pp. 190–93.

8. Note on the Theory of Heat Conduction, M. S. van Dusen. *Journal of Research*, National Bureau of Standards, Washington, D. C., vol. 4, 1930, pp. 753–56.

CONDUCTION OF HEAT IN SOLIDS (book), H. S. rslaw, J. C. Jaeger. Oxford University Press, adon, England, 1959.

TEMPERATURE FIELDS IN ELECTRICAL COILS: MERICAL SOLUTIONS, P. J. Schneider. *AIEE ansactions*, pt. I (*Communication and Electronics*), . 72, Jan. 1954, pp. 768–71.

NUMERICAL ANALYSIS OF HEAT FLOW (book), M. Dusinberre. McGraw-Hill Book Company, c., New York, N. Y., 1949.

RELAXATION METHODS IN THEORETICAL YSICS (book), R. V. Southwell. Oxford Univer- y Press, 1946.

THE NUMERICAL SOLUTION OF HEAT-CONDUC- ON PROBLEMS, H. W. Emmons. *ASME Transac- ns*, American Society of Mechanical Engineers, l. 65, 1943, pp. 607–12.

NUMERICAL MATHEMATICAL ANALYSIS (book), J. B. Scarborough. The Johns Hopkins Press, Baltimore, Md., 1959, pp. 329–40.

15. WÄRMEABGABE ÜBER DEN UMFANG EINES ANGEBLASENEN GEHEIZTEN ZYLINDERS, E. Schmidt, K. Wenner. *Forschung auf dem Gebiete des Ingenieur- wesens*, Düsseldorf, Germany, vol. 12, 1941; *NACA TM 1050*, National Advisory Committee for Aeronautics, Washington, D. C., 1943.

16. HEAT AND MASS TRANSFER (book), E. R. G. Eckert, R. M. Drake. McGraw-Hill Book Company, Inc., 1959, p. 241.

17. HEAT TRANSFER, VOL. II (book), M. Jakob. John Wiley & Sons, Inc., 1943.

18. RADIANT HEATING AND COOLING, PART I. C. O. Mackey, L. T. Wright, R. E. Clark, N. R. Gay. *Bulletin 32*, Engineering Experiment Station, Cornell University, Ithaca, N. Y., 1943.

19. RADIANT-INTERCHANGE CONFIGURATION FAC- TORS, D. C. Hamilton, W. R. Morgan. *NACA TN 2836*, National Advisory Committee for Aeronautics, 1952.

20. HEAT TRANSMISSION (book), W. H. McAdams. McGraw-Hill Book Company, Inc., 1954.

21. REVIEW OF THE THERMAL RADIATION PROPERTY VALUES FOR METALS AND OTHER MATERIALS, G. G. Gubareff, Shao-Yen Ko. *Report GR 2462-R3*, Minneapolis-Honeywell Research Center, Hopkins, Minn., 1956.

22. RADIATION ANALYSIS BY NETWORK METHOD, A. K. Oppenheim. *ASME Transactions*, American Society of Mechanical Engineers, vol. 78, 1956, pp. 725–35.

23. A SHORT METHOD FOR EVALUATING DETERMINANTS AND SOLVING SYSTEMS OF LINEAR EQUATIONS WITH REAL AND COMPLEX COEFFICIENTS, P. D. Crout. *AIEE Transactions*, vol. 60, 1941, pp. 1235–41.

Analog and Digital Computers Aid n the Design of a Numerically Controlled Positioning Servomechanism

JOHN CALLAN
ASSOCIATE MEMBER AIEE

NUMERICAL POSITIONING con- trols for machine tools have in recent ears demonstrated their ability to effect nsiderable economies in many indus- ies; hence, further applications of these ntrols to machines may be expected.

Optimum positioning performance of a achine can be obtained only if its ynamics are matched by those of the con- ol system, at least to the extent eco- omically feasible. The machine manu- acturer—although his product is de- gned on sound principles and he has kept n mind the requirements necessary for ood servomechanism performance—can eldom furnish the constants for the otion equations of the various machine otions that must be positioned by the ontrol system.

ual Purpose of Paper

Several empirical methods are known or finding a mathematical expression losely approximating the exact motions quations of a physical system. The urpose of the first part of this paper is to resent an example of one such method, mploying a digital computer to effect a solution; then, to show how a computer can aid standard servomechanism procedures to effect an optimum design, with the tedious portion of the work left to the computer.

System Type

The system described, which is characteristic of this type, saturates at a relatively small error signal. Thus, the discussion is confined to system performance with small signals. Saturation is effected at small signal levels since, to minimize positioning time, proportional action is desired only for the minimum distance from the set point necessary to achieve the required positioning accuracy and performance. Thus, the control system designer is concerned with optimizing the performance during that portion of the positioning cycle when the motion approaches its final position. The time that is required to position after the controller leaves its saturated state is called its settling time.

How to Determine Equations

Essentially, the method for determining the equations of motion makes use of the system's frequency response to effect an analytic expression for the transfer function. The plotted data from a frequency response test is the well-known Bode plot where both response of the system in decibels and phase of the system in degrees, compared with input, are plotted against the logarithm of frequency. The plot usually covers three to five decades on a semilog graph.

USE OF FREQUENCY AND TRANSIENT RESPONSE

The frequency response may be used to indicate the degree of stability of a given system and give an indication of compensating networks which will improve the performance. An analytic expression of this plot, while not necessary for determining stability, is an aid to stability improvement, and it enables the designer accurately to pinpoint the type of response the system will have to any input whatsoever.

The transfer function might also be obtained from an analysis of the transient response of the system, but this was found by trial to be less accurate than the use of the frequency response. This conclusion is supported by the work of others.[1]

Form of Transfer Function

The usual form of the transfer function is a fraction consisting of a group of first- and/or second-order factors in both the numerator and the denominator. From the slopes of the attenuation plot and the corresponding phase angles, a person with only a little practice can approximate the total order of both the numerator and denominator of the sought-after transfer function. Even if the exact order of a transfer function cannot be estimated accurately the first time, by using the method described herein, the order of the transfer function becomes apparent as the steps in this method are progressed.

Program Steps

Steps are given herewith for determining an analytic expression of a Bode plot obtained from test data:

Paper 61-38, recommended by the AIEE Industrial Control Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., January 29–February 3, 1961. Manuscript submitted October 13, 1960; made available for printing November 18, 1960.

JOHN CALLAN is with Allen-Bradley Company, Milwaukee, Wis.

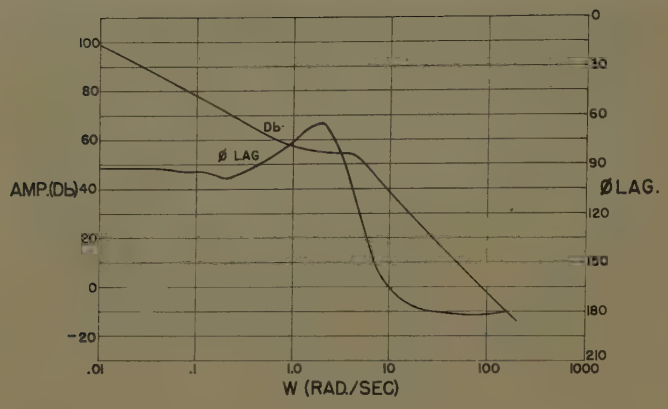


Fig. 1 (above). Bode plot of uncompensated system

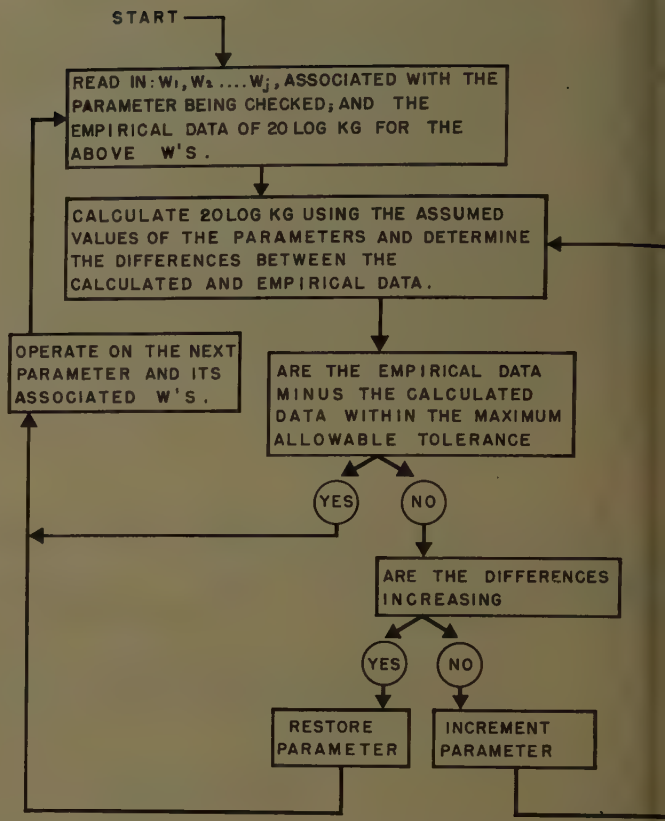


Fig. 2 (right). Computer flow diagram

1. By observation, assume the Laplace transform of the transfer function of the Bode plot to be of the form

$$KG = \frac{(s/\tau_{11}+1)\ldots(s/\tau_{1n}+1)(s^2/\omega_{21}{}^2+2\rho_{21}s/\omega_{21}+1)\ldots(s^2/\omega_{2n}{}^2+2\rho_{2n}s/\omega_{2n}+1)}{(s/\tau_{31}+1)\ldots(s/\tau_{3n}+1)(s^2/\omega_{41}{}^2+2\rho_{51}s/\omega_{41}+1)\ldots(s^2/\omega_{4n}{}^2+2\rho_{5n}s/\omega_{4n}+1)}$$

2. From the shape of the plot, estimate the $\tau_{\alpha\beta}$'s, which would be the break frequencies of the straight line approximations of the Bode plots and estimate the $\omega_{\alpha\beta}$'s and the $\rho_{\alpha\beta}$'s which are the natural frequencies and damping factors, respectively, of each second-order factor. To simplify the computer program, estimate these values for the first approximation as slightly lower than the assumed value.

3. Program the computer to calculate the values of KG for the assumed values of the parameters $\tau_{\alpha\beta}$, $\omega_{\alpha\beta}$, and $\rho_{\alpha\beta}$ at various discrete values of ω, and compare the calculated values with the observed data. Also determine by observation which parameter (i.e., $\tau_{\alpha\beta}$ or $\omega_{\alpha\beta}$ or $\rho_{\alpha\beta}$) has the most weight at each discrete ω chosen.

4. Set a reasonable limit on the maximum allowable deviation between the observed data points and the calculated values; and, if the error is outside of the reasonable limit, program the computer to increment to a larger value the magnitude of the $\tau_{\alpha\beta}$ or $\omega_{\alpha\beta}$ or $\rho_{\alpha\beta}$ which has the most weight at the point where the excessive deviation occurs.

5. By an iterative process, repeat the preceding two steps, using the incremented parameters until the deviations between the calculated data points and the empirical data points are within the specified tolerance. If the magnitude of deviation is outside of the tolerance and is increasing, the computer must be programmed to stop the iterative process and increment the parameter which caused an increased deviation to the next lower value.

6. Program the computer to store the best values of parameters as determined by step 5, and to use these parameters to recalculate and print out the values of KG selected in step 3.

7. Plot the transfer function of the computer-calculated analytic expression of the transfer function, and compare it with the plot of the empirical data. By observation, the engineer will be able to determine what changes, if any, should be made in the assumed form of the analytic expression, and steps 1 through 7 can be repeated until the data from the calculated analytic expression and the empirical data are within the tolerance specified.

These steps proceed in a manner more tedious than one would care to try without the aid of an electronic digital computer. This system, however, is an easily understood, staightforward, and speedy method of obtaining the transfer function equation, or an analytic expression from an empirical data plot.

An interesting side comment which can be made is that, before this particular method was devised, several attempts where made to adapt six different standard curve-fitting procedures to the set of data obtained from the frequency response test. These attempts were performed by mathematicians, including a representative of a well-known computer organization, who were well-versed in compute techniques; but who were not familia with servomechanism design procedures The method herein described, which wa finally used, was proposed and set up b the author, an engineer fairly well versed in servomechanism design tech niques, but only vaguely familiar wit computer technology.

This example serves to point out tha fuller use of a digital computer can b made if common practical engineerin techniques are evaluated for use on digita computer problems. Of course, this ex ample is not unidirectional, and it is im portant that the engineer should lear and apply as many digital compute techniques as is practical. The wor *digital* is purposely inserted before com puter in the foregoing statement becaus it is the author's opinion that this inter change of concepts and techniques ha

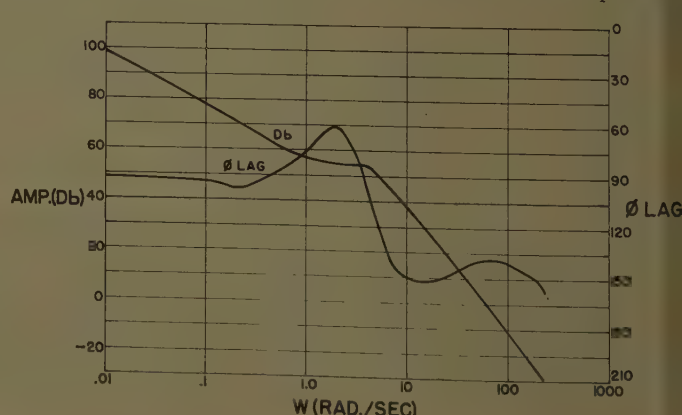
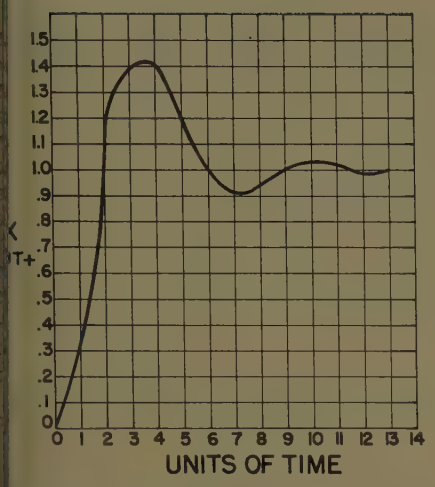


Fig. 3. Bode plot of compensated system



g. 4. Transient response of compensated system

een in practice to a great extent and r a long time in the *analog* computer eld.

lotting Attenuation Data

The seven steps previously outlined ere applied to the data taken from a ositioning table used on a layout drilling achine. The attenuation data are hown plotted in Fig. 1, in a standard anner on 5-cycle semilog graph paper. includes the transfer function of a d-c otor, driven by a thyratron controller in n open-loop configuration without any mpensating networks in the loop. The st data and the data calculated from the mputer program are a single curve in ig. 1 because the difference between the vo curves was only a fraction of a decibel t any point.

esults Obtained

A flow diagram of the computer proram is shown in Fig. 2. The results escribed were obtained after four comuter runs. The time spent by the enineer in setting up the data input for ach run was never more than 7 or less han 3 minutes, which means the transfer unction was obtained with minimum expenditure of engineering time. The results of the computer program show that the analytic expression of the data is

$$KG = \frac{(s/0.4+1)(s/1.25+1)(841.8)}{s(s/0.25+1)[(s/4.3)^2+(2)(0.45)s/4.3+1]}$$

For the first run, the computer program was set up—because of the author's faulty guess—for an expression with the form

$$KG = \frac{K(s/\tau_{11}+1)}{s(s/\tau_{21}+1)[(s/\omega_{31})^2+2\rho_{41}s/\omega_{31}+1]}$$

A short study of the first set of data from the computer program, as a result of this approximation, indicated that the actual expression had an additional factor in the numerator, and this was added for the subsequent trials. The third and fourth computations which were made were necessary only because of lack of experience with this program since they merely decreased values of the increments and of the allowable error. However, by using larger increments on the first few runs, the computer operating time was undoubtedly shortened at the expense of using slightly more engineering time. This consideration is one that probably has to be worked out at each individual facility, being a function of available computer time and cost versus engineering time and cost.

After the analytic transfer function of the inherent section of the system was calculted, a simple lead-lag network was added to enable the system to operate in a stable mode at the open-loop gain of about 60 decibels. Figs. 1 and 3, respectively, show the open-loop Bode plots without and with the stabilizing compensating network. The transfer function, then, of the stability-compensated open-loop system becomes

$$KG = \frac{(841.8)(s/0.4+1)(s/1.25+1)(s/31.2+1)}{s[(s/4.3)^2+(2)(0.45)s/4.3+1]\times(s/0.25+1)(s/156+1)}$$

and the closed-loop response can be calculated to be

$$KG = \frac{(s/0.4+1)(s/1.25+1)(s/31.2+1)}{0.00000165s^5+0.000264s^4+0.0652s^3+2.111s^2+3.333s+1}$$

The inverse Laplace transform of the system, incorporating step and ramp inputs, were obtained by appropriately modifying the expression for each type of input and solving for the output as a function of time.

Checking System Response

The calculated and empirical response of the system were checked and found to be in close agreement, but the response, shown in Fig. 4, was clearly not optimum. At this point, it must be remembered that these data and tests are for small signals which keep the system in a linear domain. To decrease the settling time of this positioning system, it seems obvious that a signal which would decrease or eliminate oscillations and overshoot is necessary.

Now the performance-optimizing that was done was only partial and, as is usually the case, it was a compromise between cost, reliability, and performance. In this case, it consisted of modifying the armature voltage with a voltage proportional to current, and using this as a damping signal. This particular part of the design was aided by an analog computer simulation of the computed system equation.

The integral of the absolute error of the system was used as the criterion for optimizing the response time. The damping signal mentioned previously was sent through a variety of frequency-sensitive networks with variable gains and a minimum value of the integrated absolute error was chosen as being this system's optimum performance. A sche-

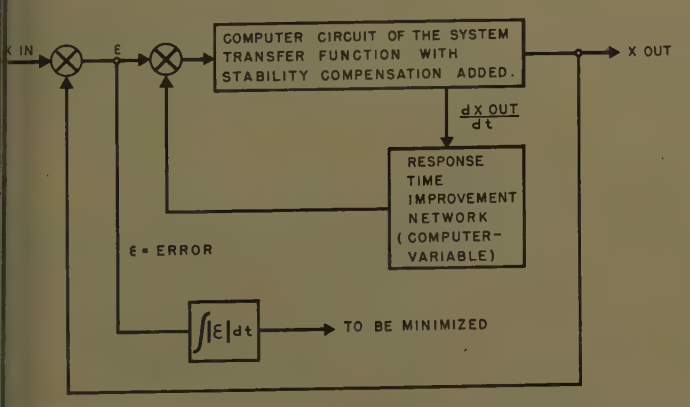


Fig. 5 (left). Block diagram of analog computer setup

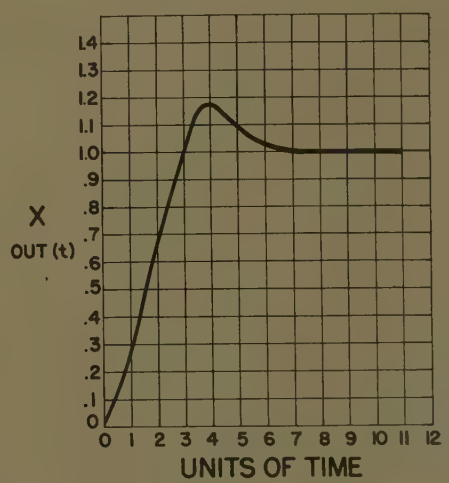


Fig. 6 (right). Transient response of optimized system

matic representation of how this was accomplished is shown in Fig. 5.

The results of this study showed that, if the damping signal described was sent through a network with a transfer function of $0.04S/(0.1S+1)$, the integral of the absolute value of the error was a minimum. The error could be made even smaller by changing the system gain by a factor of 10 and by modifying the compensation still further, but the performance improvement would be slight and the additional cost required to effect this change would not be warranted in this case. The time response of the optimized system is shown in Fig. 6.

The optimizing of a system, using an accurate mathematical representation of the system on a computer, as compared to adjusting compensating networks to improve performance in the field, is a debatable question, probably resolved best by doing both.

Conclusions

In summary, customer-engineered systems can be built more accurately, economically, and efficiently if the genera type of problem can be set up for solu tion on modern computers. For th initial problem, additional time is spent i setting up the computer program to handl a general problem of its type, but this i compensated for in customer satisfactio on the first job and the realization tha future problems in the same category wil be cut to a minimum.

Reference

1. Using Fourier Analysis in Designing, Ira Ritow. *Electrical Manufacturing*, New York N. Y., Feb. 1959, pp. 87–102.

Wide Speed Range and Torque Control of New Contactorless Precision D-C Hoist

ANSGAR HANSEN
MEMBER AIEE

JOHN H. KARLSON
ASSOCIATE MEMBER AIEE

ROBERT MIERENDORF
ASSOCIATE MEMBER AIEE

THE PERFORMANCE required of new cranes and hoists is becoming regularly more demanding; specifications now call for capabilities not previously available. Safety and dependability have always been of primary importance in hoisting work, and to these requirements are now often added a wide speed range for the rapid movement of light loads or the empty hook; quick, smooth acceleration and deceleration of all loads with generative braking to reduce service on the mechanical brake; effective torque limitation, both steady-state and transient, to limit the mechanical strains on the entire structure as well as to prevent excessive overloading; controlled slow speeds essentially independent of load; load suspension or extremely fine inching for the positioning of sensitive loads. All these characteristics and more must be furnished within the framework of a rugged, simple system requiring little or no maintenance.

Many types of hoist drive have been employed for the satisfaction of any or all of these requirements. Most are well known and it is necessary only to keep in mind that some make use of a-c motors with various schemes for the modification of the a-c motor characteristics, and that d-c motors, both series-wound and shunt-wound, are used in other drives. A characteristic of a-c motor drives is that the sizing of the motor is a function of the maximum speed and maximum load that must be handled. In the event that a 2-to-1 or 3-to-1 ratio of empty hook speed to rated hook load speed is desired, the motor name-plate horsepower that must be furnished will be in about the same proportion.

A separate high-speed motor may be furnished as an alternative. The d-c series motor has desirable characteristics for general hoisting service, but has limitations in that large currents must be switched or dissipated in heavy resistors, that operation at very light loads is uncertain, and that low-speed operation or inching at light hook loads is difficult. The d-c shunt motor is normally a device of relatively constant speed. When used in combination with an adjustable voltage supply it can provide a wide speed range with speed regulation nearly independent of load.

The Circuit

The hoist drive described in this paper makes use of a shunt-wound d-c motor in combination with an adjustable voltage control scheme which, when desired, imparts to the motor the essential characteristics of the series motor, or utilizes the constant speed properties of the shunt motor under load conditions where it is useful. This is accomplished by the use of the constant-horsepower, or field-weakening, range of the motor's characteristic to produce light hook speeds of two to four times the full-capacity speed. Typical load speed curves are shown in Fig. 1. The hyperbolic characteristics of the series motor are evident in the curves for the fourth and fifth positions of hoisting and lowering. This is operation in the motor field-weakened range. Relatively constant speed characteristics of the armature-controlled range are employed in the first three points in either direction for consistent pickup and approach performance. The curves of Fig. 1 illustrate the effect of torque (or current) limiting in that the hoisting speed falls to zero if the load exceeds some predetermined value. The load limit in the lowering direction is sufficiently higher than that in the hoisting direction to assure that the drive will lower under control any load that it is capable of lifting.

The fundamentals of the required circuitry are shown in the diagram of Fig. 2(A). A-c power is almost universally the starting point for electrically operated hoist drives; it is converted to direct current in the rotating motor-generator set. The armature of the d-c generator is connected directly to the armature of the shunt hoist motor. An amplifier numbered 2 is the exciter which drives the field of the d-c generator; amplifier 3 is the

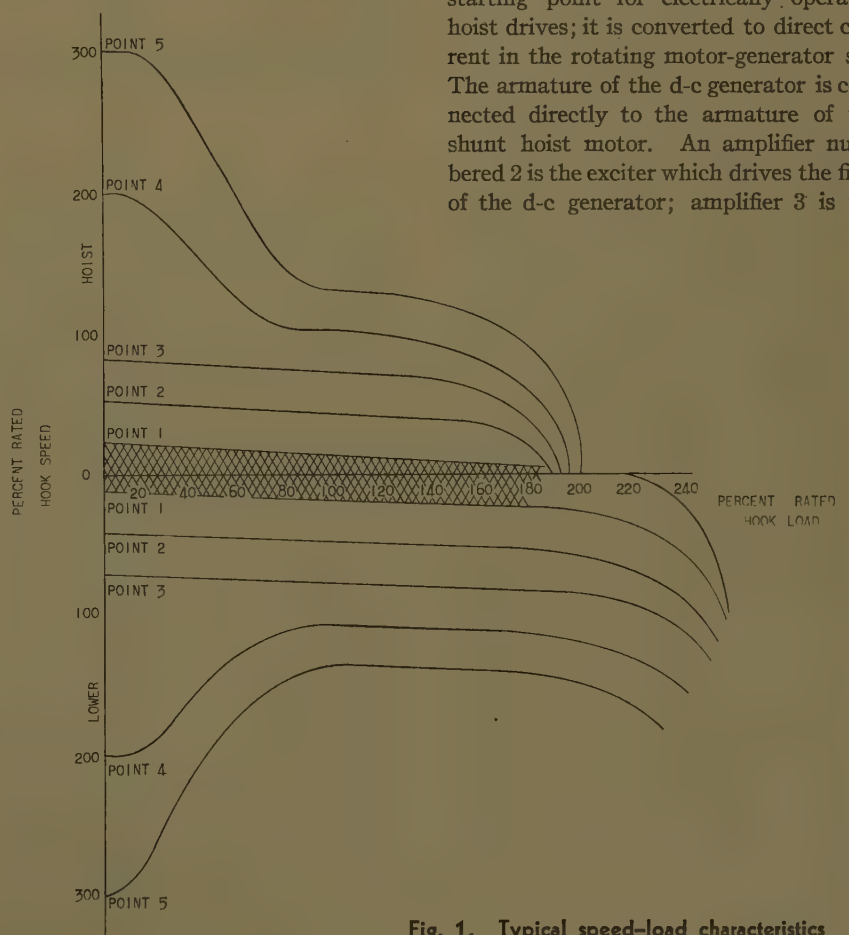


Fig. 1. Typical speed–load characteristics

Paper 60-639, recommended by the AIEE Rotating Machinery Committee of the Power Division and approved by the AIEE Technical Operations Department for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 9–11, 1960. Manuscript submitted February 23, 1960; made available for printing August 12, 1960.

ANSGAR HANSEN, JOHN H. KARLSON, and ROBERT MIERENDORF are with The Louis Allis Company, Milwaukee, Wis.

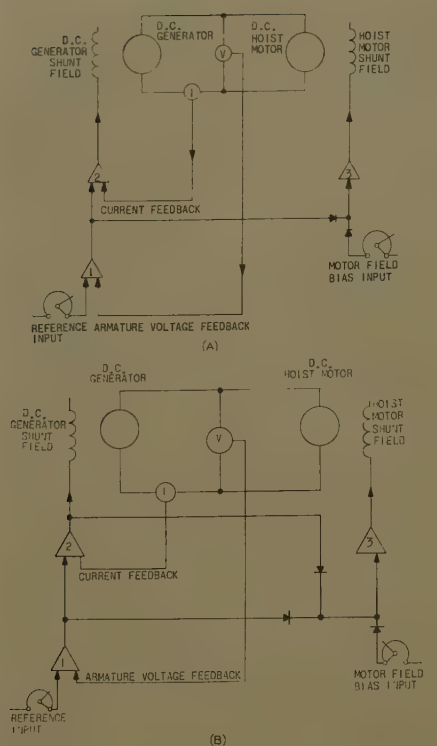


Fig. 2. A—Fundamental circuit for d-c hoist drive. B—With circuit to prevent excess regeneration on deceleration

exciter for the shunt field of the hoist motor; and amplifier 1 is the regulator.

The basic circuit is that of an armature feedback adjustable-voltage drive in which the d-c armature voltage is forced to conform in magnitude and polarity to the low-power reference signal input to amplifier 1. This arrangement in itself would result in a constant speed shunt motor characteristic. Direct armature current negative feedback is supplied to amplifier 2, thereby forcing the output of amplifier 1 to be proportional to the armature current at all loads.

If amplifier 1 were linear without limit, the current from amplifier 1 to amplifier 2 would necessarily increase in magnitude as the motor was loaded. As the negative current feedback acted to reduce the generator voltage, the slight reduction in armature voltage would force amplifier 1 to even higher outputs. Amplifier 1, however, is arranged to saturate at a suitable level and when the armature current has risen to the level which produces this saturation, the amplifier can no longer offset the effect of the negative current feedback to amplifier 2. Any further increase in load current will only serve to reduce the output voltage of the generator. The drive will stall at currents not much in excess of those at which the first amplifier saturates.

The motor shunt field amplifier 3 is biased by the control circuit to supply full motor field current or an appropriate weak field value, depending upon which characteristic it is desired to produce. The output of amplifier 1 is fed into amplifier 3 with a polarity such as always to increase the motor field current. In those positions of the master switch in which full motor field current is called for anyway, this signal is overridden by a bias signal and has no effect on the operation of the drive. At those speed points for which the exciter amplifier 3 is biased to low output, the motor field current will remain at weak field value so long as the motor armature current is not sufficiently great to drive the output of amplifier 1 to a value larger than the bias signal on amplifier 3. As the load current tends to become larger than this value, the increasing output from amplifier 1 drives the motor shunt field to higher currents, causing the hoist to operate at lower speeds.

The use of a special feedback winding on amplifier 2 carrying full armature current provides a fast, definite, current-limiting action. As a result, current overshoot (and hence peak torque) is kept down in magnitude and duration providing better protection to the electric machinery and mechanical structure. Time lags in amplifiers 1 and 2 tend to develop instability in the system when operation is in the current limit range, but this is easily eliminated by a negative rate feedback around amplifier 1.

Operation

A consideration of the events following a change in input to the drive will serve to clarify the operation of the circuit and to illuminate the desirable characteristics. Consider first the sequence of events following a change in master switch position from standstill or low-speed hoisting to a higher speed hoisting. A larger input signal to amplifier 1, unbalanced at first by any corresponding increase in armature voltage, will drive amplifier 2, and hence the generator field, to full-voltage values. As the generator voltage rises, the armature current will rise. This current will tend to limit itself through the feedback to amplifier 2 if it should exceed the desired current maximum (usually 150% to 175% of rated armature current).

If the hoist is lightly loaded, the armature voltage will come up to the value called for, the armature current will be small, and the motor will settle at the required speed. If the hoist should be heavily loaded, the armature current will continue to be large, requiring a continued large output of amplifier 1 in order to maintain the desired speed. If the position to which the master switch has been moved calls for weak field on the drive motor, the bias to amplifier 3 will have been such to produce the necessary low output from that amplifier. However, during the acceleration period, the high output of amplifier 1 will force the output of amplifier 3 to remain at full value, and the acceleration will take place at full motor field (and torque). If, when the appropriate armature voltage is reached, the load current is low the output of amplifier 3 will adjust to the low value necessary to develop a high hook speed. If the load on the hook is large and the armature current is high the continued higher output of amplifier 1 will maintain the motor field at or near its full value.

The sequence of events is much the same when the master switch is moved from some hoisting position, or from standstill, or from a low-speed lowering position to a higher speed lowering position. Assume that the point called for requires field weakening. Amplifier 1 will drive amplifier 2 and the generator field in the direction to call for the required armature voltage, now of the opposite polarity from the hoisting case. The high output of amplifier 1 keeps the exciter amplifier at full output, assuring full motor field current during the acceleration. If the load on the hook is large it will overhaul the motor and continue to accelerate the motor when the armature voltage reaches the value called for by the reference signal.

As the motor-induced voltage reaches and passes the value induced in the generator, the armature current will go to zero and reverse direction. The motor is now regenerating. If the load is large it will keep accelerating the motor until the regenerative current is large enough to counterbalance the load torque and maintain a constant lowering speed.

The action of the current limit to amplifier 2 is now in the direction to increase the induced voltage in the generator (since the regenerative current has the same direction of flow as the hoisting current). The regulator amplifier 1, however, is acting to maintain the set armature voltage and, as that voltage rises, the output of amplifier 1 will diminish to zero and reverse and increase with opposite polarity. If the load is heavy enough, this output will be sufficient so that when it is fed into amplifier 3 the motor field will be maintained at full level for slow and controlled lowering of the heavy load. In this manner, the

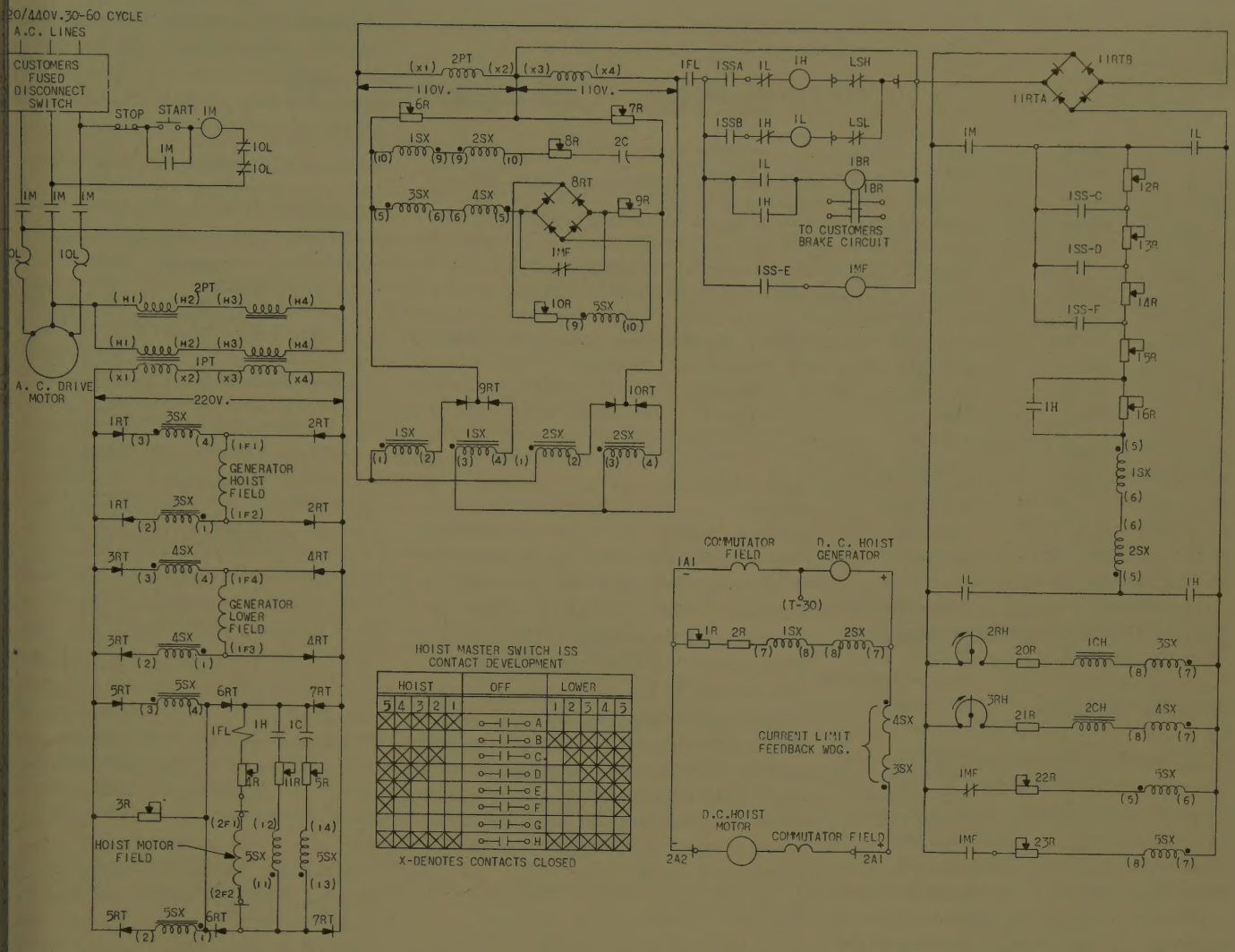


Fig. 3. Schematic diagram for hoist drive

otor field current is kept at or near full lue during all accelerations and decelera- ns, and it becomes weak only when the ad is shown to be small enough to war- nt high-speed movement. The region operation in which lowering control is complished through both the generator ld and the motor field is an inherently stable region. This, however, is over- me by manipulation of the time con- ant of the motor field amplifier 3.

The same delaying of the motor field nplifier prevents a transient dip in motor ld current during acceleration in e lowering direction when the output amplifier 1 must pass momentarily rough zero. Hence, acceleration in the wering mode and with a heavy load accomplished without the large over- oot in speed that is characteristic of me types of hoist drives.

Two factors are inherent which permit e drive to lower safely under control load larger than it can hoist. The in- rent mechanical efficiency of motor, aring, and drum arrangement is such at a larger hook load is necessary to produce the same regenerative current as is produced in hoisting a given load. Because the output of amplifier 1 must reverse and build up to saturation from a certain "negative" output when lowering an overhauling load, more armature current will be required to saturate the amplifier. If required, the circuit can be easily modified to provide a further built-in difference in limits for each direction.

In cases where a wide field range is employed (three or four to one it is possible that during deceleration from a high speed in the field range, too large a voltage may be built up in the motor armature so that saturation of the generator prevents the effective operation of the current limit. In this case a spillover circuit can be employed as shown in Fig. 2(B) to hold the motor field at a weak value until the output of amplifier 2 begins to drop off. This will provide a deceleration at constant armature current until the motor reaches about base speed; then the field will be permitted to increase as the motor comes to standstill.

A simplified schematic of the control to provide this performance is presented in Fig. 3. Here, magnetic amplifiers are shown for each of the amplifiers described above. Rotating, electronic, or semiconductor amplifiers could be used in their stead. The power circuit contains no contactors or reversing switches. The heavy current feedback keeps the circulating currents during standstill to low values. The push-pull regulator circuit provides a safe, gentle lowering of even the heaviest load if the mechanical brake should fail with the master switch in the OFF position. The reference is calling for zero armature voltage under such a condition, and the lowering speed will be less than that associated with first-point lowering.

The circuit lends itself easily to the use of stepless control. By the use of a potentiometer input and by limiting the reference voltage to that used for first-position hoisting or lowering, a continuously variable speed band between about 10% of base speed in either direction can be achieved. Of course, it is then possible to hold the load stationary on motor

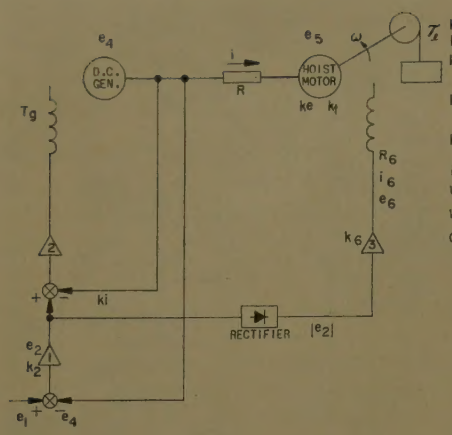


Fig. 4. Fundamental circuit used for stability analysis

torque, or to inch the hook in extremely small increments. During load suspension, armature current flows through the stationary commutator. In general, even with rated load on the hook the current will be less than rated value.

Momentary control is provided to prevent this condition from being accidentally maintained.

Stability Analysis

As previously mentioned, in the constant power range the drive is inherently unstable in the lowering mode, and special precautions must be taken to stabilize it. One method of stabilization is to slow down the response of the motor field current to change in the output signal of amplifier 1. How much damping of the motor field current is required to achieve stable operation? The necessary amount of stabilization depends on the field-weakening range. The problems involved can best be illustrated by analyzing a given drive.

Fig. 4 shows the portion of the circuit which is significant in the constant power range. The circuit constants given in the figure are assumed; they are selected in such a way that the calculations will be simplified; however, they are realistic and could very well apply to a 25-horsepower hoist. For ease of calculation, it is also assumed that the amplifier time constants are negligible in comparison to the field time constant and the mechanical time constant. This assumption can be realistic even with the use of magnetic amplifiers, if amplifiers with response approaching one half-cycle are applied.

The circuit equations for both hoisting and lowering are:

$$e_2 = k_2(e_1 - e_4) \tag{1}$$

$$e_4(1 + T_0 d/dt) = k_{34}(e_2 - ki) \tag{2}$$

$$e_4 = e_5 = Ri \tag{3}$$

$$e_5 = k_e \omega i_6 \tag{4}$$

$$\tau_l + Jd\omega/dt = k_\tau i i_6 \tag{5}$$

$$k_6 e_2 = R_6 i_6 (1 + T_m d/dt) \tag{6}$$

Because products of variables are contained in the equations the equations are nonlinear; they can be linearized by selecting certain operating points and investigating the stability in the vicinity of these points. Let one operating point for load torque be T_{l1} and the corresponding other variables be ω_1, e_{21}, e_{41}, etc. Apply a small variable change $\Delta\tau_l$ to the load, which will then assume the value $\tau_l = \tau_{l1} + \Delta\tau_l$. The corresponding values of the other variables will be $\omega = \omega_1 + \Delta\omega$, $e_2 = e_{21} + \Delta e_2$, etc. Substituting in equations 1 through 6 yields

$$\Delta e_2 = k_2 \Delta e_1 - k_2 \Delta e_4$$

$$\Delta e_4 (1 + T_0 d/dt) = k_{34} \Delta e_2 - k_{34} \Delta i$$

$$\Delta e_4 - \Delta e_5 = R \Delta i$$

$$\Delta e_5 = k_e \omega_1 \Delta k_6 + k_e i_{61} \Delta\omega + k_e \Delta i_6 \Delta\omega$$

$$\Delta\tau_l + Jd\Delta\omega/dt = k_\tau i_1 \Delta i_6 + k_\tau i_{61} \Delta i + k_\tau \Delta i \Delta i_6$$

$$k_6 \Delta e_2 = \Delta i_6 R_6 (1 + T_m d/dt)$$

Working with small variations only, the product terms of two incremental variables can be dropped so that linear equations result. Applying the Laplace transform to the equations after dropping the product terms, and letting capital letters represent the transforms, yields

$$\Delta E_2 = k_2 \Delta E_1 - k_2 \Delta E_4$$

$$\Delta E_4 (1 + T_0 s) = k_{34} \Delta E_2 - k_{34} \Delta I$$

$$\Delta E_4 = \Delta E_5 = R \Delta I$$

$$\Delta E_5 = k_e \omega_1 \Delta I_6 + k_e i_{61} \Delta\omega$$

$$\Delta\tau_l + J\Delta\omega s = k_\tau i_1 \Delta I_6 + k_\tau i_{61} \Delta I$$

$$k_6 \Delta E_2 = R_6 (1 + T_m s) \Delta I_6$$

From these equations, the response to an incremental change in any of the variable Δ-terms can be calculated. It is elected to investigate the stability by calculating the response in speed to an incremental step change in load torque $\Delta\tau_l$, assuming constant reference voltage $\Delta e_1 = 0$. Substituting $\Delta E_1 = 0$ and all the constants from Fig. 4, and leaving the motor time constant T_m in letter form yields the following:

For lowering at 2½ times base speed

$(\omega_1 = -170;\ i_1 = 100;\ i_{61} = 0.8)$

$$\frac{\Delta\omega}{\Delta\tau_l} = \frac{3.0 - (0.402T_m + 0.0006)s - 0.0006T_m s^2}{3.08 + (1.37T_m - 1.81)s + (0.405T_m + 0.0006)s^2 + 0.0006T_m s^3}$$

For lowering at 3 times base speed

$(\omega_1 = -205;\ i_1 = 100;\ i_{61} = 0.53)$

$$\frac{\Delta\omega}{\Delta\tau_l} = \frac{3.6 - (0.402T_m + 0.0006)s - 0.0006T_m s^2}{3.3 + (1.34T_m - 2.6)s + (0.28T_m + 0.00042)s^2 + 0.00042T_m s^3}$$

Clearly, the response of the hoist each of the lowering cases shown, equations 7 and 8, will be unstable unless the motor field time constant is greater than a certain amount. For a stable system the roots of the denominator of the response equations must all have negative real parts. The value of T_m required to assure stability is found to be least 1.32 seconds in the first case, and 1.95 seconds in the second.

The time constant of the motor field winding when operating in the field-weakening range is lower than these values. Thus, it is necessary to control the time constant of the motor field exciter, amplifier 3. The values calculated are borne out by the results of tests. As the field control or constant horsepower range is extended the field circuit must be made to respond more slowly in order to maintain stability. This presents no serious handicap when it is working at three or four times base speed. To keep the delay to a minimum, the higher lowering speeds can be achieved with the use of higher armature voltages necessitating a lower field range. Since no stability problem exists when hoisting in the constant power range, the field circuit damping can be eliminated in this mode; when lowering a variable or nonlinear damping may be suitable to speed the over-all response.

Summary

Careful manipulation of control functions has made it possible to generate desirable motor characteristics utilizing a basically standard type of adjustable speed drive incorporating simple and rugged circuitry. Series or shunt motor characteristics can each be employed for that part of the work which it fits best.

Power Apparatus and Systems—February 1961

Conference Papers Open for Discussion

Conference papers listed below have been accepted for AIEE Transactions and are now open for written discussion until April 26. Duplicate double-spaced typewritten copies for each discussion should be sent to Edward C. Day, Assistant Secretary for Technical Papers, American Institute of Electrical Engineers, 33 West 39th Street, New York 18, N.Y., on or before April 26.

Preprints may be purchased at 50¢ each to members; $1.00 each to nonmembers if accompanied by remittance or coupons. Please order by number and send remittance to:

AIEE Order Department
33 West 39th Street
New York 18, N. Y.

57-791 Comparison of Steel and Aluminum Subway Cars.....Bardsley
59-645 Electromagnetic Brake with Controllable Torque.....Lister
60-1020 Problems of Asymptotic Behavior and Stability.....Cesari

AIEE PUBLICATIONS

	Member Prices	Nonmember Prices: Basic Prices*†	Nonmember Prices: Extra Postage for Foreign Subscriptions
Electrical Engineering			
Official monthly publication containing articles of broad interest, technical papers, digests, and news sections: Institute Activities, Current Interest, New Products, Industrial Notes, and Trade Literature. Automatically sent to all members and enrolled students in consideration of payment of dues. (Members may not reduce the amount of their dues payment by reason of nonsubscription.) Additional subscriptions are available at the nonmember rates.		annually $12* Single copies $1.50*	$1.00
Bimonthly Publications			
Containing all officially approved technical papers collated with discussion (if any) in three broad fields of subject matter as follows:	annually	annually	
Communication and Electronics	$5.00	$8.00*	$0.75
Applications and Industry	$5.00	$8.00*	$0.75
Power Apparatus and Systems	$5.00	$8.00*	$0.75
Each member may subscribe to any one, two, or all three bimonthly publications at the rate of $5.00 each per year. A second subscription to any or all of the bimonthly publications may be obtained at the nonmember rate of $8.00 each per year.			
Single copies may be obtained when available.	$1.50 each	$1.50* each	
AIEE Transactions			
An annual volume in three parts containing all officially approved technical papers with discussions corresponding to six issues of the bimonthly publication of the same name bound in cloth with a stiff cover.	annually	annually	
Part I Communication and Electronics	$4.00	$8.00*	$0.75
Part II Applications and Industry	$4.00	$8.00*	$0.75
Part III Power Apparatus and Systems	$4.00	$8.00*	$0.75
Annual subscription to all three parts (beginning with vol. 77 for 1958).	$10.00	$20.00*	$2.25
Annual subscription to any two parts.		$15.00*	$1.50
AIEE Standards			
Listing of Standards, test codes, and reports with prices furnished on request.			
Special Publications			
Committee reports on special subjects, bibliographies, surveys, and papers and discussions of some specialized technical conferences, as announced in ELECTRICAL ENGINEERING.			

*Discount 25% of basic nonmember prices to college and public libraries. Publishers and subscription agencies 15% of basic nonmember prices. For available discounts on Standards and special publications, obtain price lists from Order Department at Headquarters.

†Foreign prices payable i
New York exchange

Send all orders to:

Order Department
American Institute of Electrical Engineers
33 West 39th Street, New York 18, N. Y.